

POLITECNICO
MILANO 1863



Silesian University of Technology

Politecnico di Milano
Department of Mechanical Engineering, Politecnico di Milano, Via La Masa 1, 20156, Milan, Italy

Deep learning-based active control of noise transmission through encapsulated structures

Author
Alkahf Aboutiman

Supervisors

Prof. Dr. Hamid Reza Karimi | Politecnico di Milano
Prof. Francesco Ripamonti | Politecnico di Milano
Prof. Marek Pawełczyk PhD, DSc | Silesian University of Technology
Dr. Stanisław Wrona | Silesian University of Technology

Year 2023-2026 Cycle 38

Abstract

Noise pollution has become a major environmental and societal issue, particularly in urban and transportation contexts, where prolonged exposure to unwanted sound can negatively affect human health, comfort, and productivity. Active noise control (ANC) offers a promising solution to this problem by using destructive interference to attenuate acoustic disturbances in real time. However, deploying effective ANC systems in real-world environments remains challenging due to complex acoustic dynamics, nonlinearities, and time-varying conditions.

This thesis investigates advanced methods for active noise control in complex vibro-acoustic environments, with a particular focus on vehicle interior applications. The main objective is to improve the robustness, adaptability, and practical performance of ANC systems.

First, a physics-based analysis of encapsulated vibro-acoustic structures is performed using the finite element method (FEM). The system is shown to exhibit non-minimum phase characteristics, which significantly limit the performance of classical adaptive ANC algorithms such as FxNLMS. To address this issue, a state-space identification framework is developed and a Kalman filter-based ANC strategy is proposed, demonstrating improved robustness and significantly enhanced noise attenuation compared to conventional approaches.

Beyond objective performance, the thesis also explores the subjective perception of ANC in encapsulated environments. Experimental studies in simulated vehicle conditions show that perceptual metrics such as annoyance and loudness are strongly influenced by both noise type and control strategy, highlighting the importance of human-centered evaluation in ANC system design.

To further improve performance in nonlinear and complex acoustic conditions, several learning-based ANC architectures are proposed. These include hybrid deep learning approaches combining convolutional neural networks with adaptive filtering, such as GFANC-THFxNLMS, which improve robustness against loudspeaker nonlinearities while maintaining computational efficiency.

Building upon these foundations, this work introduces a neural network-based ANC controller, FxMDNoiseNet, operating within a filtered-x structure. To enhance adaptability in dynamic environments, a reinforcement learning framework is integrated, leading to a deep transfer reinforcement learning (DTRL) approach. Using algorithms such as DDPG and Twin Delayed Deep Deterministic Policy Gradient (TD3), the control policy is fine-tuned online, while transfer learning from pre-trained networks accelerates convergence and improves stability. The proposed FxMDNoiseNet-DTRL framework is evaluated under realistic automotive conditions using experimentally measured impulse responses. Results demonstrate improved noise attenuation, robustness to secondary path mismatch, and strong performance under both linear time-invariant and linear time-varying conditions.

Riassunto

L'inquinamento acustico è diventato un importante problema ambientale e sociale, in particolare nei contesti urbani e dei trasporti, dove l'esposizione prolungata a suoni indesiderati può influire negativamente sulla salute umana, sul comfort e sulla produttività. Il controllo attivo del rumore (Active Noise Control, ANC) rappresenta una soluzione promettente a questo problema, poiché utilizza il principio dell'interferenza distruttiva per attenuare in tempo reale le perturbazioni acustiche. Tuttavia, l'implementazione efficace di sistemi ANC in ambienti reali rimane una sfida a causa della complessità delle dinamiche acustiche, delle non linearità e delle condizioni variabili nel tempo.

Questa tesi analizza metodi avanzati di controllo attivo del rumore in ambienti vibro-acustici complessi, con particolare attenzione alle applicazioni negli abitacoli dei veicoli. L'obiettivo principale è migliorare la robustezza, l'adattabilità e le prestazioni pratiche dei sistemi ANC.

In primo luogo, viene condotta un'analisi fisica di strutture vibro-acustiche chiuse mediante il metodo degli elementi finiti (Finite Element Method, FEM). Il sistema presenta caratteristiche di fase non minima, che limitano significativamente le prestazioni dei tradizionali algoritmi ANC adattativi, come FxNLMS. Per affrontare questo problema, viene sviluppato un approccio di identificazione in spazio di stato e viene proposta una strategia ANC basata sul filtro di Kalman, che dimostra una maggiore robustezza e un'attenuazione del rumore significativamente superiore rispetto agli approcci convenzionali.

Oltre alle prestazioni oggettive, la tesi esplora anche la percezione soggettiva dell'ANC in ambienti confinati. Studi sperimentali condotti in condizioni simulate di abitacolo automobilistico mostrano che parametri percettivi quali il fastidio percepito e la sensazione di intensità sonora sono fortemente influenzati sia dal tipo di rumore sia dalla strategia di controllo adottata, evidenziando l'importanza di una valutazione centrata sull'utente nella progettazione dei sistemi ANC.

Per migliorare ulteriormente le prestazioni in condizioni acustiche non lineari e complesse, vengono proposte diverse architetture ANC basate sull'apprendimento automatico. Tra

queste figurano approcci ibridi di deep learning che combinano reti neurali convoluzionali e filtraggio adattativo, come GFANC-THFxNLMS, in grado di aumentare la robustezza rispetto alle non linearità introdotte dagli altoparlanti mantenendo al contempo un'elevata efficienza computazionale.

Sulla base di questi risultati, il presente lavoro introduce un controllore ANC basato su rete neurale, denominato FxMDNoiseNet, operante all'interno di una struttura filtered-x. Per incrementare l'adattabilità in ambienti dinamici, viene integrato un framework di reinforcement learning, dando origine a un approccio di deep transfer reinforcement learning (DTRL). Utilizzando algoritmi quali Deep Deterministic Policy Gradient (DDPG) e Twin Delayed Deep Deterministic Policy Gradient (TD3), la politica di controllo viene ottimizzata online, mentre il transfer learning da reti pre-addestrate accelera la convergenza e migliora la stabilità del sistema.

Il framework proposto, FxMDNoiseNet-DTRL, viene infine valutato in condizioni automobilistiche realistiche utilizzando risposte impulsive misurate sperimentalmente. I risultati dimostrano un miglioramento dell'attenuazione del rumore, una maggiore robustezza rispetto agli errori di modellazione del percorso secondario e prestazioni elevate sia in condizioni lineari tempo-invarianti (LTI) sia in condizioni lineari variabili nel tempo (LTV).

Streszczenie

Zanieczyszczenie hałasem stało się istotnym problemem środowiskowym i społecznym, szczególnie w środowiskach miejskich oraz transportowych, gdzie długotrwała ekspozycja na niepożądane dźwięki może negatywnie wpływać na zdrowie człowieka, komfort oraz produktywność. Aktywna kontrola hałasu (Active Noise Control, ANC) stanowi obiecujące rozwiązanie tego problemu, wykorzystując zjawisko interferencji destrukcyjnej do tłumienia zakłóceń akustycznych w czasie rzeczywistym. Jednak skuteczne wdrożenie systemów ANC w rzeczywistych warunkach pozostaje wyzwaniem ze względu na złożoną dynamikę akustyczną, nieliniowości oraz zmienność warunków w czasie.

Niniejsza rozprawa doktorska bada zaawansowane metody aktywnej kontroli hałasu w złożonych środowiskach wibroakustycznych, ze szczególnym uwzględnieniem zastosowań we wnętrzach pojazdów. Głównym celem jest poprawa odporności, adaptacyjności oraz praktycznej skuteczności systemów ANC.

W pierwszej kolejności przeprowadzono analizę fizyczną zamkniętych struktur wibroakustycznych z wykorzystaniem metody elementów skończonych (Finite Element Method, FEM). Wykazano, że badany system charakteryzuje się właściwościami fazy nieminimalnej, które znacząco ograniczają skuteczność klasycznych adaptacyjnych algorytmów ANC, takich jak FxNLMS. W celu rozwiązania tego problemu opracowano metodę identyfikacji w przestrzeni stanów oraz zaproponowano strategię ANC opartą na filtrze Kalmana, która zapewnia większą odporność oraz znacznie lepszą redukcję hałasu w porównaniu z metodami konwencjonalnymi.

Oprócz wyników obiektywnych, rozprawa analizuje również subiektywną percepcję aktywnej kontroli hałasu w środowiskach zamkniętych. Badania eksperymentalne przeprowadzone w symulowanych warunkach wnętrza pojazdu wykazały, że takie wskaźniki percepcyjne jak uciążliwość oraz odczuwana głośność są silnie zależne zarówno od rodzaju hałasu, jak i zastosowanej strategii sterowania. Wyniki te podkreślają znaczenie oceny skoncentrowanej na użytkowniku w projektowaniu systemów ANC.

W celu dalszej poprawy skuteczności działania w warunkach nieliniowych i złożonych zaproponowano kilka architektur ANC opartych na uczeniu maszynowym. Obejmują

one hybrydowe podejścia deep learning łączące konwolucyjne sieci neuronowe z filtracją adaptacyjną, takie jak GFANC-THFxNLMS, które zwiększają odporność na nieliniowości wprowadzane przez głośniki przy jednoczesnym zachowaniu wysokiej efektywności obliczeniowej.

Na bazie tych osiągnięć w pracy przedstawiono kontroler ANC oparty na sieci neuronowej, nazwany FxMDNoiseNet, działający w strukturze filtered-x. W celu zwiększenia adaptacyjności w dynamicznych środowiskach akustycznych zintegrowano framework uczenia przez wzmacnianie, prowadząc do opracowania podejścia deep transfer reinforcement learning (DTRL). Wykorzystując algorytmy takie jak Deep Deterministic Policy Gradient (DDPG) oraz Twin Delayed Deep Deterministic Policy Gradient (TD3), polityka sterowania jest optymalizowana online, natomiast transfer learning z wykorzystaniem wcześniej wytrenowanych sieci przyspiesza proces uczenia i poprawia stabilność systemu.

Ostatecznie zaproponowany framework FxMDNoiseNet-DTRL został oceniony w realistycznych warunkach motoryzacyjnych z wykorzystaniem eksperymentalnie zmierzonych odpowiedzi impulsowych. Uzyskane wyniki wykazały poprawę tłumienia hałasu, większą odporność na niedopasowanie ścieżki wtórnej oraz wysoką skuteczność zarówno w warunkach liniowych stacjonarnych w czasie (LTI), jak i liniowych zmiennych w czasie (LTV).

Acknowledgements

This dissertation was carried out as a double-degree doctorate under the agreement between the Silesian University of Technology, Poland, and Politecnico di Milano, Italy. A part of the work was supported by the project IN-NOVA: Active reduction of noise transmitted into and from enclosures through encapsulated structures, which has received funding from the European Union's Horizon Europe program under the Marie Sklodowska-Curie grant agreement no. 101073037, coordinated by the Silesian University of Technology.

I would like to express my deepest gratitude to my supervisors Prof. Hamid Reza Karimi and Prof. Francesco Ripamonti from Politecnico di Milano, Italy for their support during my PhD and their precious advice for this project. I would also like to thank Prof. Marek Pawełczyk and Prof. Stanisław Wrona from the Silesian University of Technology, Poland for their guidance, encouragement, and insightful feedback throughout this work.

I also would like to thank Dr. Chao Liang, Dr. Zulfi Rachman and Dr. Khaled Said Ahmed Maamoun for the collaboration in the research conducted together and in the writting of articles. I am thankful to master students that I have co-supervised, Maxime Zaman and Reza Shams for their involvement and resilience in their research.

I am grateful to Prof. Jian Kang, Prof. Francesco Aletta and Dr. Tin Oberman from University College London, UK, for hosting me during my secondment at their esteemed institute, and to Dr. Rolf Schirmacher and Dr. Jan Kirchhof for their cooperation and hospitality during my secondment at Müller-BBM Active Sound Technology GmbH, Germany. My appreciation further extends to all members of the International Supervisory Council of the IN-NOVA project.

I am also thankful to colleagues from Politecnico di Milano, University College London, Müller-BBM and the INNOVA projects for their collaboration and valuable discussions.

I am deeply thankful to my family and friends for their unwavering support, love, and patience, which made this journey possible.

Contents

Abstract **i**

Riassunto **iii**

Streszczenie **v**

Acknowledgements **vii**

Contents **ix**

List of Figures **xiii**

List of Tables **xvii**

List of Abbreviations **xix**

List of Symbols **xxi**

1 Introduction **1**

1.1 Background: Active Noise Control and Its Limitations 1

1.2 Active Noise Control in Encapsulated Structures 2

1.3 Learning-Based Active Noise Control 4

1.4 Thesis Outline 6

2 Active noise control on non-minimum phase system 9

2.1 Definition of Non-Minimum Phase Systems 9

2.2 Finite Element Model Development 10

2.2.1 Geometry and model definition 10

2.2.2 Physics modeling 11

2.2.3 Computation of the control paths 13

2.2.4 Encapsulated structure with non-minimum phase characteristics . . 14

2.3 Active Noise Control on Non-Minimum Phase Systems 15

2.3.1 FxNLMS on non-minimum phase systems 15

2.3.2 Kalman filter-based active noise control 17

2.3.3 Computational complexity 18

2.4 Results and discussion 19

2.4.1 Noise reduction performance 19

2.4.2 Discussion 21

3 Objective and subjective evaluation of hybrid deep learning based active noise control through encapsulated structure 23

3.1 SFANC-NLMS algorithm through encapsulated structure 23

3.1.1 Experimental set-up 23

3.1.2 SFANC-NLMS algorithm 25

3.1.2.1 Selective Fixed-Filter Active Noise Control (SFANC) . . . 26

3.1.2.2 CNN-Based SFANC-NLMS Framework 27

3.1.2.3 Hybrid SFANC-NLMS Approach 29

3.1.3 Results 30

3.2 Subjective perception of Noise Reduction Achieved Using a Hybrid Deep Learning-Based Active Noise Control Algorithm 31

3.2.1 Experimental Procedure 31

3.2.2 Data Analysis 33

3.2.3 Discussion . . . 34
3.2.3.1 Effect of ANC algorithms on perceived annoyance . . . 34
3.2.3.2 Effect of sound level on perceived annoyance under different ANC conditions . . . 35
3.2.3.3 Interaction between noise sources, ANC condition, and sound levels . . . 36
3.3 Hybrid Generative Deep Learning-Based ANC for Encapsulated Structures with Openings under Nonlinearities . . . 37
3.3.1 Environment of control . . . 38
3.3.2 The GFANC-THFxNLMS algorithm . . . 39
3.3.2.1 THFxNLMS adaptive control stage . . . 39
3.3.2.2 GFANC Method . . . 42
3.3.2.3 GFANC-THFxNLMS Approach . . . 44
3.3.2.4 Computational complexity . . . 46
3.3.3 Results . . . 46

4 Deep transfer reinforcement learning based active noise control in vehicle interior 51

4.1 Environment of control . . . 51
4.2 Deep learning based active noise control . . . 52
4.2.1 Filtered-x Modified DNoiseNet Algorithm . . . 52
4.2.2 MDNoiseNet architecture . . . 54
4.2.3 Training strategy . . . 55
4.3 Deep transfer reinforcement learning active noise control . . . 56
4.3.1 Twin Delayed Deep Deterministic Policy Gradient . . . 57
4.3.1.1 Implementation of the algorithm . . . 57
4.3.1.2 Training Strategy . . . 59
4.4 Results . . . 60

5 Conclusions **63**

5.1 Basic framework and motivations of appended papers 63

5.2 Summary of paper A 64

5.3 Summary of paper B 64

5.4 Summary of paper C 65

5.5 Summary of paper D 65

5.6 Summary of paper E 66

5.7 Future work 67

Bibliography **69**

Attached Papers **75**

List of publications **153**

List of Figures

2.1 Schematic of the Noise Box with dimensions used for the finite element model [30]. 11

2.2 Pole-zero diagrams in the s-domain for the primary path (a), secondary path (b), and reference path (c). 15

2.3 Comparison of adaptive control strategies for active noise control in NMP systems: (a) FxNLMS algorithm, (b) Kalman filter-based approach. 16

2.4 Time-domain and time-frequency representation of the control performance. (a) FxNLMS residual error signal. (b) Kalman filter residual error signal. (c) Spectrogram comparison between uncontrolled and controlled cases. . . 20

3.1 Noise Box: (a) laboratory prototype; (b) schematic representation of the vibroacoustic enclosure [29]. 24

3.2 Images of the experimental set-up: (a) outside view and (b) inside view. . . 25

3.3 Block diagram of the hybrid SFANC-NLMS algorithm. For each frame, the 1D CNN running on a co-processor selects the most appropriate pre-trained control filter. The NLMS algorithm then refines and updates the coefficients of the selected filter in real time at the sampling rate. 25

3.4 Block diagram of the selective fixed-filter active noise control method inspired from [16]. 26

3.5 Frequency-domain comparison of NLMS, SFANC, and SFANC-NLMS for motorbike noise. 30

3.6 Images of the experimental set-up from inside (a) and outside (b) [46]. . . 32

3.7 Mean perceived annoyance (PAY) and standard deviations for different ANC conditions across noise levels. 35

3.8 Circumplex representation of perceived affective quality (PAQ) across sound levels. ISOPL (pleasantness) is plotted against ISOEV (eventfulness) for 55, 65, and 72 dB(A). 36

3.9 Mean perceived loudness (PLN) and standard deviations across noise levels for different ANC conditions and noise types. 37

3.10 Overview of the experimental rig showing a vibrating steel plate with a central circular opening of radius 0.05 m. Microphones (Mic 1, Mic 2, and Mic 3) are positioned at different locations within the enclosure. Mic 1 is used as the error microphone, while Mic 2 and Mic 3 serve as observation points. The primary noise source is placed outside the enclosure, and the secondary loudspeaker is positioned behind the opening. 39

3.11 Block diagram of the hybrid GFANC-THFxNLMS algorithm. f_{LS} models the loudspeaker nonlinearity. A 1D CNN running on a co-processor selects the most suitable pre-trained control filter, which is then refined online using the THFxNLMS algorithm. The arrow from the "Generated control filter" block to $W(z)$ denotes a one-time initialization: at the start of each processed signal, $W(z)$ is initialized with the CNN-generated filter, after which it is continuously updated sample-by-sample by the THFxNLMS adaptation rule, without further frame-wise resets from the CNN [60]. . . . 40

3.12 Scaled error function (SEF) for different values of η. 48

3.13 PSD comparison 49

3.14 Noise reduction per second 49

3.15 Comparison of spectral performance and temporal noise reduction under varying SNR conditions. 49

4.1 Experimental vehicle used for the measurements: (a) interior view and (b) exterior view. 52

4.2 FxMDNoisenet for feedback ANC. The network takes the estimated disturbance $\hat{d}(n)$ as input and generates the control signal $y(n)$, which is filtered through the secondary path $S(n)$ to produce the anti-noise $a(n)$. The residual $e(n)$ is obtained by subtracting $a(n)$ from the disturbance $d(n)$. . . 53

4.3 Architecture of the proposed MDNoiseNet. The network is composed of 1D causal, pointwise, and atrous convolutions to extract effective features from raw time-domain acoustic signals. The hyperparameters of each layer are specified in the form (input channels, output channels, kernel size, dilation). For layers without dilation, the dilation factor is denoted as d when applicable. 55

4.4 FxMDNoiseNet-DTRL control framework. The agent observes the state s_n (estimated disturbance) and generates the control action a_n. The residual error is used to compute the reward r_n, which guides policy learning. . . . 56

4.5 Actor and critic parameters updates. 59

4.6 Comparison of (a) NMSE over time and (b) PSD for FxNLMS (green), FxMDNoiseNet (blue), and FxMDNoiseNet-DTRL (red) under LTV conditions with combined engine and road noise. The estimated secondary path corresponds to the empty vehicle, while the true path corresponds to the vehicle with two occupants. 61

4.7 Spectrogram comparison of primary noise and error signals obtained with FxNLMS, FxMDNoiseNet, and FxMDNoiseNet-DTRL under LTV conditions. 62

List of Tables

2.1 Physical properties of the aluminum plate 12

2.2 Computational complexity and noise reduction performance of FxNLMS and Kalman filter algorithms 19

3.1 Frequency ranges associated with the pre-trained control filters. 27

3.2 Architecture of the proposed 1D CNN model. 28

3.3 Comparison between the proposed CNN architecture and existing one-dimensional neural network models. 29

3.4 Noise reduction performance for different noise types (dB). 31

3.5 Architecture of the 1D CNN Model. 44

3.6 Computational complexity of the proposed 1D CNN model. 44

3.7 Step-by-step description of the GFANC-THFxNLMS algorithm with corresponding equations. 45

3.8 Complexity analysis of the GFANC-THFxNLMS algorithm. 47

4.1 TD3 training hyperparameters used for the ANC task. 60

List of Abbreviations

ANC Active Noise Control

CNN Convolutional Neural Network

CRN Convolutional Recurrent Network

DDPG Deep Deterministic Policy Gradient

DNoisenet Deep NoiseNet

DTRL Deep Transfer Reinforcement Learning

FEM Finite Element Method

FFT Fast Fourier Transform

FRF Frequency Response Function

Fx Filtered-x structure

FxLMS Filtered-x Least Mean Square

FxMDNoiseNet Filtered-x Modified DNoiseNet

FxNLMS Filtered-x Normalized Least Mean Square

GFANC Generative Filtered-x Active Noise Control

LTI Linear Time-Invariant

LTV Linear Time-Varying

MDNoiseNet Modified DNoiseNet

MIMO Multiple Input Multiple Output

NLMS Normalized Least Mean Square

NMP Non-Minimum Phase

RL Reinforcement Learning

RNN Recurrent Neural Network

SFANC Selective Filtered-x Active Noise Control

SISO Single Input Single Output

TD3 Twin Delayed Deep Deterministic Policy Gradient

List of Symbols

ρ — Air density (kg/m^3)
c — Speed of sound (m/s)
p — Acoustic pressure (Pa)
$e(n)$ — Residual error signal at time step n
$d(n)$ — Primary disturbance signal
$a(n)$ — Anti-noise signal generated by the secondary path
$y(n)$ — Control signal generated by the ANC controller
$\hat{d}(n)$ — Estimated disturbance signal
$s(n)$ — Secondary path impulse response at time n
$\hat{s(n)}$ — Estimated secondary path impulse response
$S(z)$ — Secondary path transfer function
$P(z)$ — Primary path transfer function
$R(z)$ — Reference path transfer function
$S(n)$ — Secondary path impulse response (time domain)
$\hat{S}(n)$ — Estimated secondary path impulse response
$*$ — Linear convolution operator
$\mathbf{w}$ — Adaptive filter coefficient vector
μ — Step size of FxNLMS algorithm
N — Number of filter coefficients
L — Length of input signal window (MDNoiseNet input)
γ — Discount factor in reinforcement learning
π_ϕ — Deterministic policy (actor network)
ϕ — Actor network parameters
θ_i — Critic network parameters (TD3)
$Q_\theta(s, a)$ — Action-value function (critic)
$r(n)$ — Reward signal in reinforcement learning
s_n — State at time step n
a_n — Action at time step n
$N_{\text{ex},n}$ — Exploration noise in TD3

τ — Soft update coefficient (Polyak averaging)
d_n — Episode termination indicator
$\mathcal{D}$ — Replay buffer
λ_{out} — Output energy regularization weight
λ_Q — Critic regularization coefficient
λ_π — Actor regularization coefficient
ω — Angular frequency (rad/s)
f_{dom} — Dominant frequency of residual signal
f_{ex} — Exploration frequency
ρ_n — Exploration gain
φ_n — Random phase in exploration signal
$K(n)$ — Kalman gain matrix
$P(n)$ — Error covariance matrix
Q — Process noise covariance (Kalman filter)
Z_R — Radiated acoustic impedance
η — Structural loss factor
Q_v — Volume velocity of monopole source (m^3/s)
$f_{\max}$ — Maximum frequency of analysis (Hz)
$\lambda_{\max}$ — Maximum wavelength criterion for meshing

1 | Introduction

1.1. Background: Active Noise Control and Its Limitations

Active Noise Control (ANC) is a well-established technique for attenuating unwanted acoustic and vibration signals based on the principle of destructive interference. The fundamental idea consists in generating a secondary signal, commonly referred to as "anti-noise", which has the same amplitude as the primary disturbance but opposite phase, resulting in partial or complete cancellation of the sound field [1, 2]. This principle has been extensively studied in the literature and forms the basis of many modern control strategies for acoustic noise reduction.

The effectiveness of ANC systems relies on the accurate estimation of both the amplitude and phase of the disturbance signal at a given spatial location. While amplitude variations are often relatively slow, the phase of acoustic waves is highly sensitive to spatial position and continuously evolves over time due to propagation effects [3, 4]. This inherent property of wave propagation makes ANC a challenging problem, particularly in real-world environments where acoustic fields are dynamic and influenced by reflections, structural interactions, and environmental variations.

Over the past decades, ANC has been successfully applied in a wide range of engineering domains, including automotive cabins, aircraft interiors, industrial machinery, household appliances, and personal audio devices. In these applications, the main objective is to improve acoustic comfort and reduce noise exposure for end users. In enclosed environments such as vehicle cabins or machinery enclosures, passive noise control techniques—based on sound absorption, insulation, and damping materials—have traditionally been employed [5, 6]. However, these methods become increasingly inefficient at low frequencies, where achieving significant attenuation requires bulky and heavy structures, which is often incompatible with industrial design constraints.

To overcome these limitations, ANC has emerged as a promising alternative, particu-

larly for low-frequency noise attenuation. In automotive applications, for instance, interior noise is typically dominated by engine-induced tonal components and broadband road–tire interactions, with most of the energy concentrated below 500 Hz [7, 8]. These low-frequency disturbances are further amplified by acoustic resonances within the cabin, making them difficult to mitigate using passive approaches alone.

Classical ANC systems are predominantly based on adaptive filtering techniques, such as the Filtered-x Least Mean Square (FxLMS) algorithm and its variants. These methods are widely used due to their simplicity, robustness, and relatively low computational cost. However, their performance strongly depends on the accuracy of system modeling, particularly in the representation of acoustic paths between the control source and error sensors. In practical applications, this requirement introduces significant challenges when system conditions vary or when modeling uncertainties are present.

Moreover, classical ANC algorithms are derived under the assumption of linear acoustic behavior. In practice, real systems often exhibit deviations from this assumption due to hardware imperfections, actuator limitations, and nonlinear response characteristics. These effects can degrade the performance of adaptive algorithms, leading to reduced convergence speed, steady-state errors, and limited robustness in real-world implementations. These limitations highlight the need for more advanced ANC strategies capable of handling complex acoustic environments while maintaining real-time implementation feasibility.

1.2. Active Noise Control in Encapsulated Structures

Encapsulated structures are widely encountered in engineering applications where acoustic isolation between an internal cavity and the external environment is required. Typical examples include automotive cabins, aircraft fuselages, railway compartments, and industrial enclosures. In these systems, noise sources are often located both inside and outside the enclosure, and the resulting acoustic field is shaped by complex interactions between structural boundaries, internal reflections, and source excitation mechanisms [9, 10].

From an acoustic standpoint, enclosed environments are characterized by strong modal behavior, particularly at low frequencies, where the wavelength becomes comparable to the characteristic dimensions of the cavity. In this regime, standing waves and resonance phenomena dominate the sound field, leading to spatially non-uniform pressure distributions. A single-channel ANC system, relying on one secondary source and one error microphone, can only minimize noise locally, within a small region around the error sensor commonly referred to as the “zone of quiet”; outside this region, attenuation may be

negligible or the noise may even be locally amplified. These effects make global noise reduction challenging and motivate the use of spatially distributed (multichannel) control strategies, which employ multiple secondary sources and error microphones, typically formulated as a multiple-input multiple-output (MIMO) adaptive filtering problem, to widen the controlled region and shape the sound field over a larger portion of the cavity [11, 12].

In addition to modal behavior, practical encapsulated structures often include openings or apertures required for ventilation, cooling, structural integration, or user interaction. While necessary for functional reasons, these openings significantly modify the acoustic response of the system by allowing partial sound transmission between the interior and exterior domains. As a result, the sound field inside the enclosure becomes highly sensitive to boundary conditions and external disturbances [13, 14].

In automotive applications, active noise control is primarily motivated by the need to improve passenger comfort by reducing interior cabin noise, which remains a key factor in perceived vehicle quality. The acoustic environment inside a vehicle is dominated by multiple noise sources, among which engine-related noise plays a particularly significant role. Engine noise is inherently periodic and strongly correlated with the crankshaft rotation, generating tonal components and harmonics that propagate through the vehicle structure via multiple mechanical paths. These structure-borne vibrations are transmitted to the cabin panels and subsequently radiated as airborne sound, resulting in a complex and highly coupled vibro-acoustic field [10, 15]. Unlike stationary or isolated noise sources, engine noise exhibits strong non-stationary behavior due to continuous variations in engine speed, load conditions, and driving dynamics. These variations induce continuous shifts in the spectral content of the acoustic field, making real-time tracking and adaptation particularly challenging for active noise control systems. Furthermore, the propagation of engine-induced vibrations through the vehicle body introduces multiple coupled transmission paths, which are highly dependent on structural configuration and operating conditions. The presence of cabin resonances further amplifies specific frequency components, particularly in the low-frequency range below 500 Hz, where traditional passive treatments are least effective. These characteristics make engine noise one of the most challenging disturbances to mitigate using active control techniques, as effective attenuation requires both accurate modeling of the transmission paths and fast adaptation to time-varying excitation conditions. As a result, automotive ANC systems must be designed to operate under highly dynamic and uncertain acoustic environments while maintaining stability and robustness in real-time implementations.

1.3. Learning-Based Active Noise Control

Recent advances in active noise control have led to the emergence of learning-based frameworks that aim to improve adaptability and robustness in complex acoustic environments. These methods extend classical adaptive filtering strategies by introducing data-driven components capable of selecting, tuning, or directly generating control signals based on learned representations of the noise field. In this context, learning-based ANC can be viewed as an evolution of traditional approaches such as the Filtered-x Least Mean Square (FxLMS) algorithm, with the objective of improving performance in non-stationary and uncertain environments.

A first category of learning-based ANC methods is represented by Selective Fixed-Filter Active Noise Control (SFANC). In this framework, the selection mechanism is typically implemented as a lightweight classifier, often a compact neural network, that maps short frames of the incoming noise signal to the index of the most suitable pre-trained filter within the bank. In this framework, a set of pre-trained control filters is designed offline for different noise conditions, and a selection mechanism is used to choose the most appropriate filter in real time. This approach reduces computational complexity during operation and provides a structured way to handle variability in noise characteristics. However, its performance strongly depends on the quality and diversity of the pre-trained filter bank, which limits its adaptability to unseen acoustic conditions [16].

To improve the flexibility of SFANC, hybrid configurations such as SFANC-FxNLMS have been introduced. In this approach, a learning-based module is used to select or combine filters, while the FxNLMS algorithm performs real-time adaptation of the control coefficients. This combination allows the system to benefit from both data-driven decision-making and adaptive filtering stability. As a result, SFANC-FxNLMS improves robustness in non-stationary environments while maintaining the stability properties of classical adaptive ANC methods [16, 17].

More recently, generative fixed-filter frameworks such as Generative Fixed-filter Active Noise Control (GFANC) have been proposed to further enhance flexibility. In GFANC, a single broadband control filter is decomposed into multiple sub-filters, which are linearly combined using weights determined by a learning-based model, typically a neural network that predicts, frame by frame, the combination weights directly from features extracted from the incoming noise signal. This structure reduces the need for multiple independent filter training processes and allows more compact representations of the control solution. GFANC-based methods are particularly suitable for embedded applications, where computational efficiency and memory constraints are critical [18, 19].

In parallel, deep learning-based ANC approaches have emerged as a direct alternative to classical adaptive filtering. These methods aim to learn a direct mapping between acoustic observations and control actions using deep neural networks. Unlike traditional algorithms, which rely on explicit error minimization rules, Deep ANC models, based on Convolutional Recurrent Neural Networks (CRN), are trained to approximate optimal control behavior from data. This paradigm enables improved handling of nonlinear and time-varying acoustic phenomena, especially in scenarios where system modeling is difficult or incomplete [20, 21]. However, despite their strong performance, these models often involve a large number of parameters, which can become a critical limitation for real-time ANC applications due to computational constraints, latency requirements, and limited embedded hardware resources.

Within this context, several deep architectures have been proposed for different ANC-related tasks. For instance, CsNNet has been introduced for complex sound separation and noise suppression in real-world environments. It relies on deep convolutional feature extraction to improve robustness under highly variable acoustic conditions, demonstrating strong generalization capabilities beyond controlled laboratory setups [21]. This class of models highlights the ability of deep learning to handle realistic and unpredictable noise scenarios.

Among recent developments, convolutional-recurrent architectures such as DNoiseNet have attracted significant attention. DNoiseNet combines convolutional neural networks (CNNs) for spectral feature extraction with recurrent neural networks (RNNs), to capture temporal dependencies in non-stationary noise signals. This hybrid design enables simultaneous modeling of short-term spectral structure and long-term temporal evolution, leading to improved performance in dynamic acoustic environments [22]. This is a lightweight architecture, making it well suited for real-time applications.

Finally, reinforcement learning has recently been integrated into deep learning-based active noise control frameworks as a promising strategy to further enhance adaptability and robustness. In these approaches, the ANC problem is formulated as a sequential decision-making task, where an agent learns an optimal control policy through direct interaction with the acoustic environment. This formulation enables online adaptation of control parameters and improved performance under time-varying and uncertain acoustic conditions. For instance, GFANC-RL combines fixed-filter decomposition with a reinforcement learning algorithm based on Soft Actor-Critic (SAC). This approach enables continuous adaptation of control weights without requiring explicit labeled data or precise secondary path identification, while preserving the structural efficiency of the GFANC framework [23]. Similarly, a narrowband ANC study using Deep Deterministic Policy Gra-

dient (DDPG) applies deep reinforcement learning without relying on a secondary path model, demonstrating robust control performance and stability under path variations in real-time experiments [24]. In addition, other studies have demonstrated that deep reinforcement learning can achieve stable noise attenuation even in the absence of accurate system models, highlighting its potential for model-free ANC design [25]. Early investigations based on Q-learning have also shown that reinforcement learning can maintain robust control performance under sudden variations of the secondary path, although these results remain largely limited to simplified or simulation-based environments [26, 27]. Despite these promising results, reinforcement learning-based ANC methods are still at an early stage of development, particularly for realistic automotive applications. Challenges related to training stability, sample efficiency, and real-time implementation remain open research problems. Moreover, most existing studies are validated in controlled or simulated environments, which limits their direct applicability to complex in-vehicle acoustic conditions. These limitations highlight the need for more efficient, stable, and physically consistent learning frameworks capable of reliable deployment in real-world ANC systems.

1.4. Thesis Outline

To address the challenges associated with active noise control (ANC) in complex acoustic environments, particularly in encapsulated structures and vehicle interiors, this thesis investigates a set of learning-based control strategies ranging from classical adaptive filtering to deep learning and reinforcement learning approaches. The main objective is to improve robustness, adaptability, and perceptual performance of ANC systems under realistic and time-varying conditions.

The main contributions of this thesis are as follows:

- Development of a comprehensive modeling and identification framework for active noise control in non-minimum phase (NMP) systems. In addition, Kalman filter-based approaches are investigated to enhance estimation accuracy and robustness under dynamic and uncertain acoustic conditions, improving performance over classical adaptive filtering methods.
- Development of a novel hybrid GFANC-THFxNLMS framework for encapsulated structures with acoustic openings, combining GFANC-based filter decomposition with a nonlinearity-aware adaptive stage (THFxNLMS) whose compression parameter is adapted online from the instantaneous error power, jointly addressing filter generation and transducer-induced nonlinearity within a single architecture.

- Objective and subjective evaluation of learning-based ANC algorithms, including perceptual studies assessing the impact of different control strategies on perceived annoyance and acoustic comfort under various noise types and sound pressure levels.
- Design of deep learning-based ANC architectures capable of handling nonlinear and time-varying acoustic systems, while maintaining computational efficiency through lightweight neural network structures.
- Design of FxMDNoiseNet, a new lightweight ANC architecture, integrated into a deep transfer reinforcement learning framework (FxMDNoiseNet-DTRL), where a DDPG/TD3 policy is initialised from pre-trained FxMDNoiseNet weights, envelope-scaled exploration, and exposure to two measured secondary paths for robustness to real path mismatch.

The structure of the thesis is organized as follows.

Chapter 2 focuses on active noise control in non-minimum phase (NMP) systems. This chapter presents the physical and mathematical modeling of the considered acoustic environment, with particular emphasis on the challenges introduced by NMP dynamics. A system identification framework is developed to estimate the reference, primary, and secondary acoustic paths. In addition, Kalman filter-based approaches are investigated to improve estimation accuracy and control performance under time-varying conditions.

Chapter 3 addresses hybrid deep learning-based ANC in encapsulated structures. This chapter introduces ANC frameworks combining classical adaptive filters with deep learning components, including filter decomposition and selection strategies inspired by SFANC and GFANC approaches. Both objective performance metrics and subjective perceptual evaluations are used to assess the effectiveness of the proposed methods.

Chapter 4 focuses on deep transfer reinforcement learning for active noise control in vehicle interiors. This chapter introduces a reinforcement learning-based ANC framework designed for nonlinear, time-varying, and non-stationary acoustic environments. A transfer learning strategy is used to improve training efficiency and convergence. The proposed method is validated in realistic automotive-like scenarios.

Chapter 5 provides a concluding discussion of the research carried out in this thesis. It summarizes the main findings, highlights the key contributions, and outlines potential directions for future work in learning-based active noise control.

2 | Active noise control on non-minimum phase system

Active noise control (ANC) is an effective approach for reducing low-frequency noise in enclosed environments where passive solutions are limited by size and weight constraints. However, its performance strongly depends on the acoustic dynamics of the system. In practical applications such as vehicle cabins and encapsulated structures, the control paths may exhibit non-minimum phase (NMP) behavior, which introduces phase delays and causality constraints that degrade the performance of conventional adaptive algorithms such as FxNLMS.

This chapter addresses ANC in NMP systems using an encapsulated vibro-acoustic model representative of automotive environments. First, the NMP characteristics are introduced through the finite element model and system identification of the control paths. Then, two control strategies are studied: the classical FxNLMS algorithm and a Kalman filter-based approach. Their computational complexity is also discussed. Finally, their noise reduction performance is compared through numerical simulations.

2.1. Definition of Non-Minimum Phase Systems

A linear time-invariant (LTI) system is classified as a minimum phase system when all its zeros lie inside the stability region, i.e., inside the unit circle in the discrete domain or in the left-half complex plane in the continuous domain. Conversely, a system is referred to as NMP if it contains at least one zero outside this stability region. Mathematically, for a discrete transfer function

$$H(z) = \frac{B(z)}{A(z)}, \tag{2.1}$$

the system is NMP if at least one zero satisfies

$$|z_i| > 1. \tag{2.2}$$

In the continuous domain, a system is considered NMP if it has zeros in the right-half plane, i.e.,

$$\Re(s_i) > 0. \tag{2.3}$$

Unlike unstable systems, NMP systems remain stable but exhibit an inherent phase lag that cannot be removed without introducing instability. This property leads to fundamental limitations in inversion-based control strategies, which are commonly used in active noise control applications. In particular, the presence of unstable zeros imposes several practical limitations on ANC design. First, the ideal inversion-based control filter $W(z) = P(z)/S(z)$ becomes unstable whenever $S(z)$ is NMP, since inverting its unstable zeros produces unstable poles. Second, achieving a stable and causal approximation of this filter necessarily introduces phase distortion and delay, which cannot be fully compensated without violating causality. Third, these effects fundamentally bound the achievable attenuation and convergence speed of gradient-based adaptive algorithms such as FxNLMS, particularly around the frequencies associated with the unstable zeros.

2.2. Finite Element Model Development

To investigate the behavior of NMP vibro-acoustic systems, a finite element model (FEM) representative of an enclosed structure is developed. The objective is to reproduce the main structural-acoustic interactions occurring in vehicle-like cavities while maintaining a computational framework suitable for control-oriented analysis.

The considered system consists of a rigid-walled acoustic cavity coupled with a flexible aluminium plate acting as the main radiating surface. This configuration is commonly used as a simplified representation of enclosed environments such as vehicle cabins, where the coupling between structural vibration and internal pressure fields plays a dominant role in low-frequency noise transmission.

2.2.1. Geometry and model definition

The finite element model is based on a vibro-acoustic configuration representative of a vehicle-like enclosed structure, commonly referred to as a Noise Box. The main objective is to reproduce the essential acoustic and structural dynamics of enclosed cavities while

limiting computational cost by avoiding the modeling of the external acoustic domain.

The system consists of an internal acoustic cavity coupled with a flexible rectangular aluminum plate acting as the main radiating surface. Only the air domain inside the cavity and the vibrating plate are explicitly modeled. The surrounding external field is neglected in order to reduce computational effort and focus on internal sound propagation and fluid–structure interaction mechanisms.

The acoustic cavity is modeled as a pentagonal prism with a local cavity at the location of the vibrating plate. Its overall dimensions are $0.825\,\mathrm{m} \times 1.018\,\mathrm{m} \times 0.645\,\mathrm{m}$. A rectangular aluminum plate, with dimensions $a \times b = 0.825\,\mathrm{m} \times 1.018\,\mathrm{m}$, is mounted at the cavity opening and defines the flexible boundary of the system.

The model is implemented using COMSOL Multiphysics® [28], following a modeling approach consistent with the vibro-acoustic plate-cavity characterization presented in [29], which allows fully coupled vibro-acoustic simulations within a unified finite element framework.

The overall geometry of the Noise Box used in this study is illustrated in Figure 2.1.

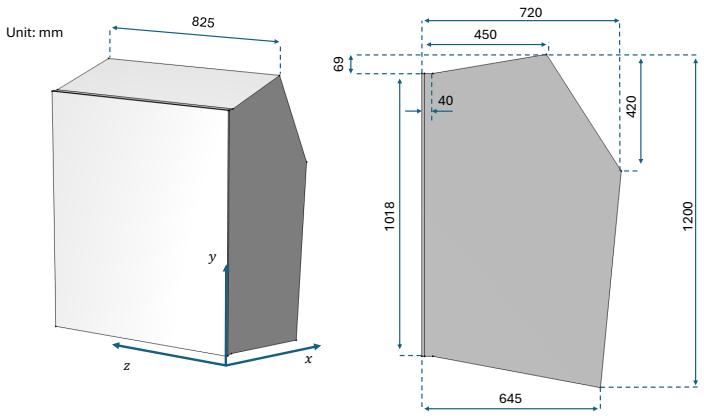


Figure 2.1: Schematic of the Noise Box with dimensions used for the finite element model [30].

The aluminium plate properties used in the model are summarized in Table 2.1.

2.2.2. Physics modeling

The structural domain is modeled using a 3D Shell formulation based on the Mindlin–Reissner theory with MITC (Mixed Interpolation of Tensorial Components) elements. This formulation allows accurate representation of both thin and moderately thick shells. However, given the very small thickness-to-span ratio of the plate, $h/a = 0.0048 \ll 1/20$, the

Table 2.1: Physical properties of the aluminum plate

Property	**Value**
Thickness, t	$0.0004\,\mathrm{m}$
Density, ρ	$2700\,\mathrm{kg\,m^{-3}}$
Young's modulus, E	$70 \times 10^9\,\mathrm{Pa}$
Poisson's ratio, ν	0.33
Structural loss factor, η	0.003

structure is in the thin-shell regime. Consequently, the Kirchhoff–Love theory is the most appropriate approximation for the present study [31–33]. The shell model provides three degrees of freedom per node in the x, y, and z directions. Structural damping is introduced through an isotropic loss factor $\eta = 0.003$, selected based on typical values for aluminum plates [34]. The plate edges are assumed to be perfectly clamped to the cavity boundary.

The acoustic domain is modeled in the frequency domain assuming linear wave propagation in air. The air properties are defined as $\rho_0 = 1.2043\,\mathrm{kg/m^3}$ and $c = 343.2\,\mathrm{m/s}$ at 20°C. All cavity walls, except the vibrating plate, are assumed to be acoustically rigid, i.e., $\overline{v}_n = 0$.

To account for acoustic loading and reverberant effects, an impedance boundary condition is introduced using $Z_0 = 100 \times \rho_0 \times c$. This impedance is used to define the radiated impedance of the vibrating structure:

$$Z_R = \frac{\hat{p}(\omega)}{i\omega\hat{x}(\omega)} \tag{2.4}$$

where $\hat{p}$ is the acoustic pressure and $\hat{x}$ is the normal displacement of the plate surface. The impedance is complex and can be expressed as:

$$Z_R = R_R + iX_R \tag{2.5}$$

where R_R represents the radiation resistance (damping effect) and X_R represents the radiation reactance (added mass effect), both influencing the frequency response amplitude and phase [35, 36].

The plate is discretized using quadratic shell elements, with a maximum element size of $0.024\,\mathrm{m}$, while the acoustic (air) domain is discretized using quadratic tetrahedral

elements, with a maximum element size $\lambda_{\max}$ defined by

$$\lambda_{\max} = \frac{c}{N \cdot f_{\max}}, \tag{2.6}$$

where $f_{\max} = 1000\,\text{Hz}$ is the maximum frequency of interest and N denotes the number of mesh elements per acoustic wavelength; a value of $N = 7$ to 10 is adopted, in accordance with standard finite element meshing guidelines for acoustic wave propagation problems [37].

2.2.3. Computation of the control paths

The finite element model is used to compute the frequency response functions (FRFs) associated with the control paths required for active noise control (ANC). In this study, three main transfer paths are considered: the reference path, the primary path, and the secondary path. These FRFs define the dynamic behavior of the vibro-acoustic system and serve as the basis for system identification and controller design.

The overall control configuration is illustrated in Figure 2.3, where the different paths used in the ANC algorithm are highlighted. To compute these FRFs, the system is excited using an incident plane wave applied to the vibrating aluminum plate. The acoustic excitation is defined through the combination of incident and reflected pressure fields, ensuring a realistic representation of an external disturbance acting on the structure.

The incident and reflected acoustic pressures are given by:

$$p_{\text{inc}} = Ae^{i\phi}e^{-i(k_x x + k_y y + k_z z)}e^{i\Phi} \tag{2.7}$$

$$p_{\text{refl}} = Ae^{i\phi}e^{-i(-k_x x + k_y y + k_z z)}e^{i\Phi} \tag{2.8}$$

with the total pressure defined as:

$$p_{\text{total}} = p_{\text{inc}} + p_{\text{refl}}. \tag{2.9}$$

The wave numbers are defined as $k_x = k\sin(\theta)\sin(\phi)$, $k_y = k\sin(\theta)\cos(\phi)$, and $k_z = k\cos(\theta)$, where θ and ϕ describe the direction of incidence.

A Single-Input Single-Output (SISO) formulation is adopted for each transfer path. The reference path is defined as the ratio between the incident acoustic pressure and the

normal acceleration of the plate. The primary path corresponds to the transfer between the incident pressure and the pressure measured inside the cavity at the error microphone location. The secondary path models the effect of the control source, represented by a monopole actuator with volume velocity $Q = 1\,\mathrm{m}^3/\mathrm{s}$. The secondary path transfer function is then obtained as the ratio between the monopole volume velocity and the resulting acoustic pressure at the error microphone location. The FRFs are computed over the frequency range from 10 Hz to 1000 Hz with a step of 1 Hz.

The computed frequency response functions are subsequently used to derive reduced-order models suitable for control design. This step is essential to obtain a compact representation of the vibro-acoustic dynamics while preserving the main physical characteristics of the system over the frequency band of interest.

To this end, each transfer function is approximated in a rational form using a residue–pole representation:

$$H(s) = \sum_n \frac{C_n}{s - A_n} + D \tag{2.10}$$

where $s = i\omega$, C_n are the residues, A_n are the poles, and D is the direct transmission term. This technique is described in detail in [38, 39]. Once the rational approximation is obtained, a continuous-time state-space. Finally, the obtained state-space models are discretized using a Zero-Order Hold (ZOH) method with a sampling frequency of 8000 Hz, allowing their direct implementation in digital control algorithms.

2.2.4. Encapsulated structure with non-minimum phase characteristics

To further investigate the dynamic properties of the system, the identified continuous-time transfer functions are analyzed in the complex s-plane. In particular, the pole–zero distributions are extracted from the rational models obtained through the AAA-based identification procedure. This representation provides direct insight into the stability and minimum-phase characteristics of each control path. Figures 2.2 show the pole–zero maps of the reference, primary, and secondary paths, respectively.

The presence of transmission zeros located in the right half of the complex plane confirms the NMP nature of the system. This characteristic is a direct consequence of the coupled vibro-acoustic dynamics, where structural resonances and acoustic wave propagation introduce non-causal phase behavior in the identified input–output relationships.

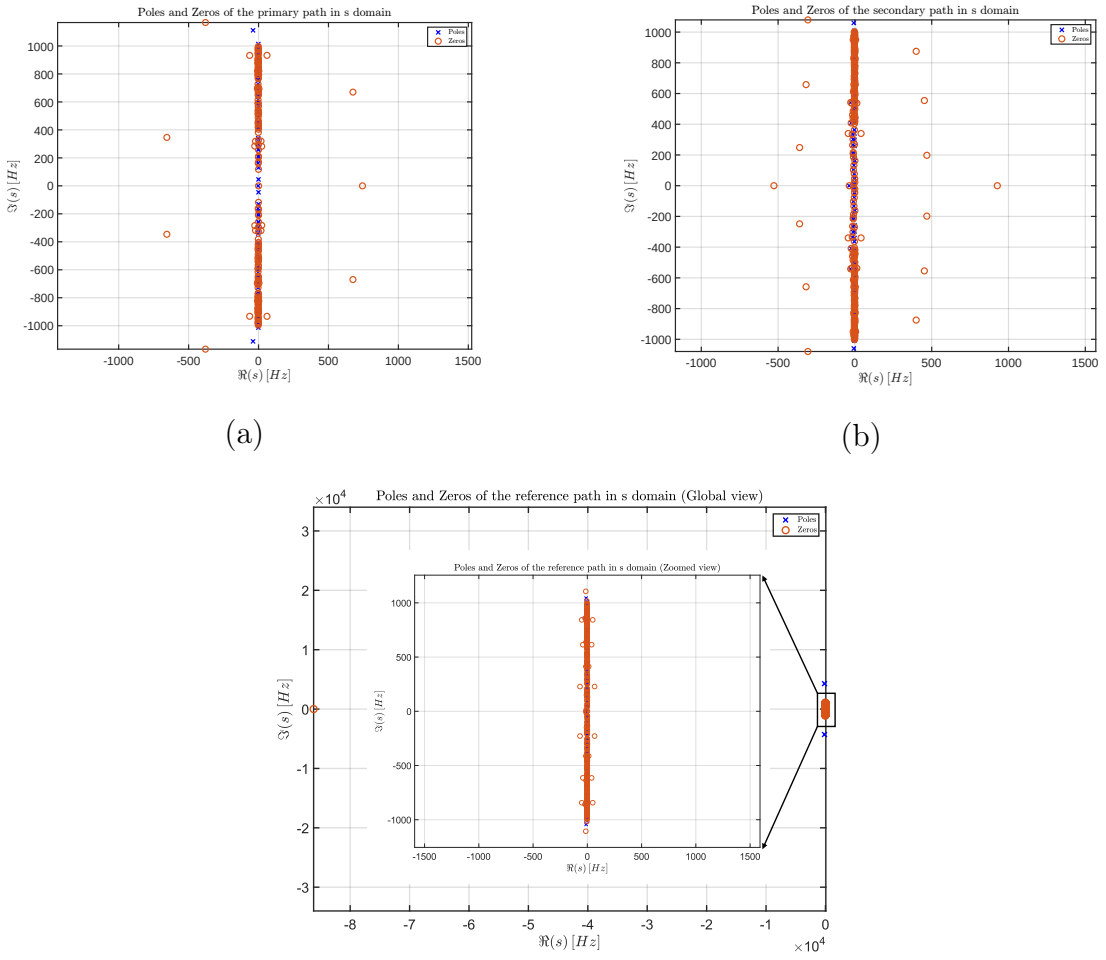


Figure 2.2: Pole-zero diagrams in the s-domain for the primary path (a), secondary path (b), and reference path (c).

2.3. Active Noise Control on Non-Minimum Phase Systems

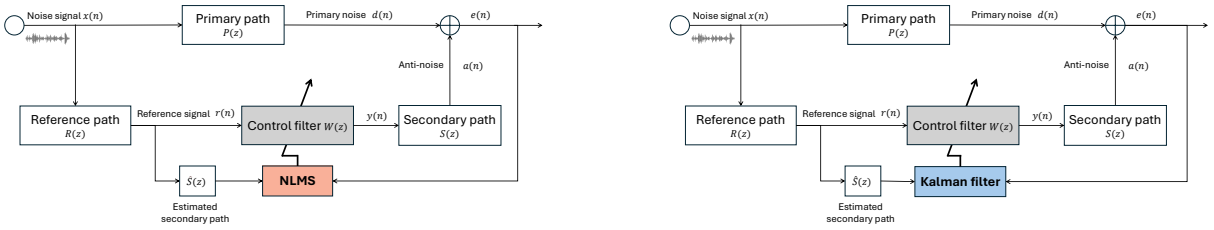
2.3.1. FxNLMS on non-minimum phase systems

The Filtered-x Normalized Least Mean Square (FxNLMS) algorithm is widely used in feedforward active noise control (ANC) systems to compensate for the effect of the secondary path. However, its performance becomes significantly degraded when the acoustic environment exhibits NMP behavior.

In an ideal ANC configuration, the optimal control filter is given by $W(z) = P(z)/S(z)$, where $P(z)$ and $S(z)$ denote the primary and secondary paths, respectively. In NMP

systems, the secondary path may contain zeros outside the unit circle, and their inversion leads to unstable poles and additional phase distortions. This fundamentally limits the achievable performance of adaptive algorithms relying on secondary path inversion.

The FxNLMS algorithm attempts to mitigate this issue by introducing an estimate of the secondary path $\hat{S}(z)$, which is used to filter the reference signal before adaptation. The resulting control architecture is illustrated in Figure 2.3a.



(a) FxNLMS algorithm structure.

(b) Kalman filter-based ANC structure.

Figure 2.3: Comparison of adaptive control strategies for active noise control in NMP systems: (a) FxNLMS algorithm, (b) Kalman filter-based approach.

Let $x(n)$ denote the reference signal vector and $r(n)$ the filtered reference signal obtained after propagation through the reference path. The controler output is expressed as follows:

$$y(n) = w^T(n)r(n) \tag{2.11}$$

Where $w(n)$ denotes the control filter coefficient vector and $w^T(n)$ its transpose. The residual error signal at the error microphone is then given by:

$$e(n) = d(n) - s(n) * y(n) \tag{2.12}$$

where $d(n)$ is the disturbance and $s(n)$ is the impulse response of the secondary path. The FxNLMS update rule is derived using a stochastic gradient descent on the instantaneous cost function $J(n) = e^2(n)$, leading to:

$$w(n+1) = w(n) + \mu e(n) \frac{r'^*(n)}{r'^H(n)r'(n) + \epsilon} \tag{2.13}$$

where $r'(n) = \hat{s}(n) * r(n)$ is the filtered reference signal and ϵ is a small regularization constant ensuring numerical stability.

In NMP systems, the presence of unstable zeros in $S(z)$ introduces significant phase distortion in the filtered reference signal $r'(n)$, as well as delays that are difficult to compensate. As a result, the adaptation process becomes slower and may suffer from performance degradation or even instability in strongly NMP configurations.

2.3.2. Kalman filter-based active noise control

To overcome the limitations of gradient-based adaptive algorithms in NMP environments, a Kalman filter-based formulation can be adopted. Unlike FxNLMS, which relies on instantaneous gradient estimation, the Kalman approach explicitly models the evolution of the control filter as a stochastic state-space system. Figure 2.3b present the diagram of the algorithm.

At convergence, the optimal control filter satisfies:

$$w(n+1) = w(n) = w_o \tag{2.14}$$

The measured disturbance can be rewritten as:

$$d(n) = r'^T(n)w_o + e_o(n) \tag{2.15}$$

where $r'(n)$ is the filtered reference signal and $e_o(n)$ represents the residual noise.

This formulation naturally leads to a state-space representation where the filter coefficients are treated as the system state. The Kalman filter recursion is then used to estimate $w(n)$ recursively.

The prediction step is given by:

$$\hat{w}(n) = w(n-1) \tag{2.16}$$

The Kalman gain is computed as:

$$K(n) = P_{\text{pred}}(n)r'(n)\left[r'^T(n)P_{\text{pred}}(n)r'(n) + q(n)\right]^{-1} \tag{2.17}$$

The state update equation becomes:

$$\hat{w}(n+1) = \hat{w}(n) + K(n)\left[d(n) - r'^T(n)\hat{w}(n)\right] \tag{2.18}$$

Finally, the error covariance matrix is updated as:

$$P(n) = (I - K(n)r'^{T}(n))P_{\text{pred}}(n) \tag{2.19}$$

Where $P_{\text{pred}}(n)$ is the predicted auto-correlation matrix of the state error at time n. Compared to FxNLMS, the Kalman filter provides a principled way to handle uncertainty in both the model and the measurements. The adaptive gain $K(n)$ automatically balances trust between the prediction and the observation, which is particularly beneficial in NMP systems where phase distortions and model mismatch are significant.

As a result, the Kalman-based approach generally exhibits improved robustness and smoother convergence behavior in complex vibro-acoustic environments, although at the cost of increased computational complexity compared to FxNLMS. This is consistent with prior studies in the ANC literature: Lopes and Piedade [40] first introduced the Kalman filter as an alternative to LMS-based adaptive algorithms, reporting improved convergence behavior under model uncertainty; van Ophem and Berkhoff [41] extended the approach to multi-channel ANC, showing improved robustness to secondary-path variations compared to conventional adaptive filtering; and Fabry et al. [42] proposed a reduced-complexity Kalman filter formulation specifically to make this improved robustness practical for real-time implementation, confirming that the computational overhead can be mitigated without fully sacrificing performance.

2.3.3. Computational complexity

This section presents a comparative analysis of the computational complexity of the FxNLMS and Kalman filter-based active noise control algorithms. The objective is to highlight the trade-off between computational burden and noise reduction performance, particularly in the context of NMP systems.

The FxNLMS algorithm remains one of the most widely used approaches in real-time ANC applications due to its low computational cost. However, its performance can be limited in complex vibro-acoustic environments, especially when phase distortions introduced by the secondary path degrade the adaptation process. Let N denote the number of coefficients of the adaptive control filter. For the considered implementation, the FxNLMS algorithm requires:

$$4N + 1 \quad \text{multiplications}, \qquad 4N + 3 \quad \text{additions}$$

where N is the number of adaptive filter coefficients. Despite its simplicity, its noise reduction capability is limited in the considered configuration.

In contrast, the Kalman filter introduces a significantly higher computational burden due to matrix operations involved in the state estimation and covariance update. The number of floating-point operations required is:

$$N^3 + 2N^2 + 4N + 1 \quad \text{multiplications}, \qquad N^3 + 3N^2 + 3N - 1 \quad \text{additions}$$

Although computationally more expensive, the Kalman filter provides substantially improved performance in terms of noise attenuation. This is mainly due to its ability to explicitly model uncertainty and recursively update the control filter based on both prediction and measurement information, which is particularly beneficial in NMP systems.

Table 2.2: Computational complexity and noise reduction performance of FxNLMS and Kalman filter algorithms

Algorithm	**FLOPs Multiplication**	**FLOPs Addition**	**Noise Reduction (dB)**
FxNLMS	$4N + 1$	$4N + 3$	2.30
Kalman Filter	$N^3 + 2N^2 + 4N + 1$	$N^3 + 3N^2 + 3N - 1$	17.10

2.4. Results and discussion

The performance of the proposed Active Noise Control (ANC) strategies is evaluated for the identified NMP vibro-acoustic system. The comparison focuses on the FxNLMS and Kalman filter algorithms in terms of transient behavior, steady-state attenuation, and frequency-domain performance under broadband excitation.

The analysis is carried out in a control configuration where the primary disturbance is generated by a filtered stochastic noise signal. This allows the assessment of robustness under realistic excitation conditions and highlights the influence of the NMP characteristics of the system on control performance.

2.4.1. Noise reduction performance

In the present study, the disturbance signal is modeled as a zero-mean Gaussian white noise with unit variance, shaped using a band-limiting filter in order to restrict its spectral content to the frequency range of interest (20–450 Hz). The filter is designed with a passband edge at 400 Hz and a stopband at 450 Hz, ensuring a smooth attenuation

of high-frequency components. The maximum passband ripple and minimum stopband attenuation are set to 0.1 dB and 80 dB, respectively.

All simulations are performed using a sampling frequency of 8 kHz, consistent with the discrete-time state-space representation of the identified model. The FxNLMS algorithm is implemented with a filter length of 512 taps and a step size of $\mu = 4.2 \times 10^{-4}$. For the Kalman filter, the observation noise variance is set to $q(n) = 0.5$. Each simulation is conducted over a duration of 5 seconds.

Figure 2.4 illustrates the time-domain behavior of the residual error signals obtained with both control strategies.

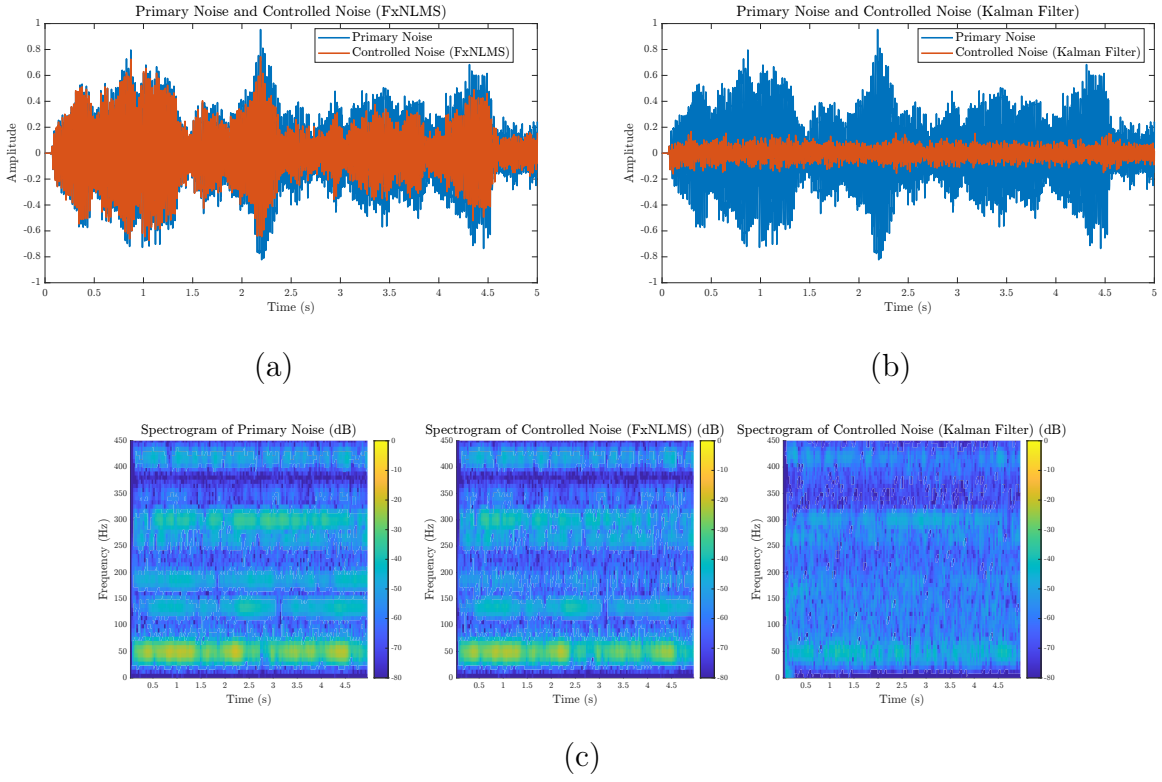


Figure 2.4: Time-domain and time-frequency representation of the control performance. (a) FxNLMS residual error signal. (b) Kalman filter residual error signal. (c) Spectrogram comparison between uncontrolled and controlled cases.

The results clearly show that the Kalman filter achieves a more stable and consistent attenuation of the disturbance across time. In contrast, the FxNLMS algorithm exhibits slower convergence and higher residual fluctuations, which are amplified by the NMP behavior of the system, particularly through phase distortion induced by the secondary path estimation.

In the spectrogram representation, the Kalman-based controller provides a more uniform

attenuation over the entire frequency band, whereas the FxNLMS shows localized regions of reduced performance, especially near resonant modes of the primary path.

This performance difference can be directly linked to the NMP nature of the system. As observed in the frequency response analysis, the primary path exhibits strong resonances and anti-resonances, leading to the classical waterbed effect. This phenomenon implies that attenuation in one frequency band is necessarily accompanied by amplification in another band.

In this context, the Kalman filter demonstrates better adaptability to spectral redistribution effects induced by the NMP behavior. It is particularly effective in attenuating resonant peaks while maintaining overall stability of the control loop. However, at anti-resonance frequencies, the residual noise may locally exceed the uncontrolled case, which is a direct consequence of the fundamental limitations imposed by the system dynamics.

2.4.2. Discussion

The obtained results highlight the strong influence of the NMP characteristics of the vibro-acoustic system on the performance of active noise control algorithms. Although the system remains stable from a physical standpoint, its phase behavior introduces significant challenges in the design of effective control strategies, particularly in feedforward ANC configurations.

The FxNLMS algorithm, while computationally efficient, shows limited performance in this context. Its reliance on an accurate secondary path model makes it particularly sensitive to phase distortions and delays introduced by the NMP dynamics. As a result, convergence is slower and residual noise attenuation remains limited, especially in frequency regions dominated by strong resonances and anti-resonances. These limitations are further amplified at higher frequencies, where shorter wavelengths increase the sensitivity of the system to modeling errors and spatial discretization effects.

In contrast, the Kalman filter demonstrates improved robustness with respect to modeling uncertainties and system dynamics. By explicitly incorporating state estimation and uncertainty propagation, it provides a more adaptive framework for tracking the optimal control solution. This leads to improved attenuation performance and more stable behavior across the considered frequency range. However, despite these advantages, performance degradation is still observed at higher frequencies. This limitation is primarily attributed to the increased complexity of acoustic wave propagation in enclosed cavities, where multiple reflections and spatial variations reduce the effectiveness of global control strategies.

Another important aspect concerns computational complexity. Although the Kalman filter outperforms FxNLMS in terms of noise reduction, this improvement comes at the cost of significantly higher computational demand. This trade-off is critical for real-time implementation, particularly in embedded ANC systems where computational resources are limited.

3 | Objective and subjective evaluation of hybrid deep learning based active noise control through encapsulated structure

3.1. SFANC-NLMS algorithm through encapsulated structure

3.1.1. Experimental set-up

The experimental investigation is carried out using the Noise Box, a laboratory-scale vibroacoustic enclosure designed to emulate the acoustic environment of a vehicle interior. As illustrated in Figure 3.1, the structure consists of a rigid concrete enclosure with a single flexible aluminium panel acting as the only compliant boundary of the cavity. The remaining walls are highly rigid, ensuring strong acoustic confinement inside the enclosure.

The aluminium panel, with a thickness of 4 mm, is clamped to the concrete frame using a steel mounting structure and a dense screw pattern, ensuring repeatable boundary conditions and a well-defined structural-acoustic coupling. This configuration creates a coupled vibroacoustic system suitable for active noise control investigations in enclosed environments [29].

The experimental objective is to characterise the vibroacoustic control paths required for active noise control within the Noise Box. These include the reference path, primary path, and secondary path, which define the dynamic relationships between excitation sources, structural response, and internal acoustic pressure.

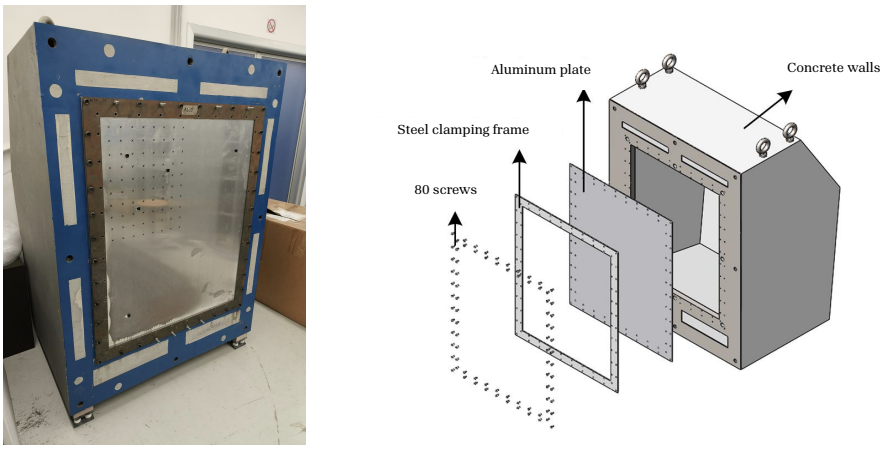


Figure 3.1: Noise Box: (a) laboratory prototype; (b) schematic representation of the vibroacoustic enclosure [29].

To simulate external acoustic disturbances, a loudspeaker is positioned outside the enclosure. The resulting structural response of the aluminium panel is measured using an accelerometer, which serves as the reference sensor. The accelerometer is mounted at a non-nodal location on the panel in order to ensure reliable modal excitation and avoid measurement cancellation effects.

Inside the cavity, a microphone is used as the error sensor to capture the residual acoustic pressure. Although multiple microphones are available within the enclosure, only one is used in this study to maintain a single-channel control configuration.

For the secondary path identification, an internal loudspeaker is installed inside the cavity and used as the control actuator. Both external and internal loudspeakers are driven using broadband random excitation signals to ensure sufficient spectral content for system identification over the frequency range of interest.

The measurement chain consists of a PCB 130E20 microphone with a sensitivity of 37.37 mV/Pa, calibrated using a Brüel & Kjær 4231 calibrator at 1 kHz and 94 dB. Structural vibrations are measured using a PCB 333B30 accelerometer with a sensitivity of 99.6 mV/g (10.16 mV/m/s^2).

All sensors are connected to a PCB Piezotronics 483C signal conditioner and subsequently interfaced with a National Instruments NI cDAQ-9178 data acquisition system. Excitation signals are generated from a computer and simultaneously transmitted to the external and internal loudspeakers through the acquisition system, ensuring synchronized actuation and measurement.



Figure 3.2: Images of the experimental set-up: (a) outside view and (b) inside view.

3.1.2. SFANC-NLMS algorithm

This section presents the hybrid SFANC-NLMS algorithm used in this study and its implementation in the experimental vibroacoustic environment introduced in Section 3.1.1. The proposed approach combines the Selective Fixed-Filter Active Noise Control (SFANC) framework with the normalized least mean square (NLMS) adaptive algorithm in order to improve the robustness of active noise control in a non-minimum phase (NMP) system. The overall architecture of the hybrid SFANC-NLMS algorithm is illustrated in Figure 3.3.

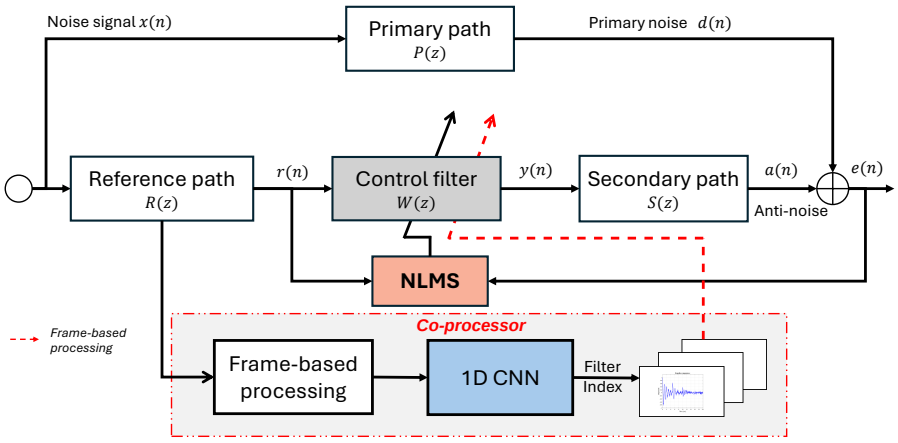


Figure 3.3: Block diagram of the hybrid SFANC-NLMS algorithm. For each frame, the 1D CNN running on a co-processor selects the most appropriate pre-trained control filter. The NLMS algorithm then refines and updates the coefficients of the selected filter in real time at the sampling rate.

3.1.2.1. Selective Fixed-Filter Active Noise Control (SFANC)

The Selective Fixed-Filter Active Noise Control (SFANC) method is a hybrid ANC strategy based on a database of pre-trained control filters and an intelligent filter selection mechanism [16]. Unlike conventional adaptive algorithms, SFANC separates the filter optimization stage from the real-time control stage. The complete framework is illustrated in Figure 3.4.

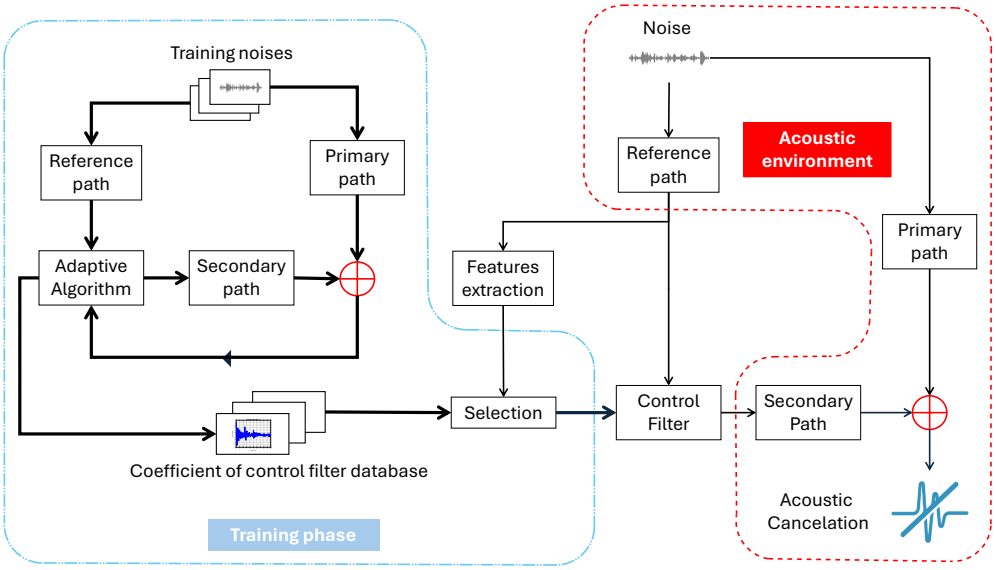


Figure 3.4: Block diagram of the selective fixed-filter active noise control method inspired from [16].

The SFANC method is divided into two main stages:

1. **Offline filter pre-training**

   During the offline stage, several control filters are optimized in advance using synthesized noise disturbances with different spectral contents. For each of the seven frequency ranges listed in Table 3.1, a broadband white noise signal restricted to the corresponding band is synthesized and injected as the disturbance in the simulated Noise Box environment. The NLMS algorithm, with a control filter length of $L = 4096$ taps and a fixed step size $\mu = 0.0001$, is then run offline until convergence for each of the seven bands independently. The converged filter coefficients $\mathbf{w}_i$, $i = 1, \ldots, 7$, are stored as fixed FIR filters in the control filter database, indexed by C_i as summarized in Table 3.1. Each stored filter therefore encodes the optimal control solution for its associated frequency range and remains unchanged during online operation, until selected and further refined by the adaptive stage described below.

The different frequency ranges used for filter generation are summarized in Table 3.1.

Control filter (C_i)	**Frequency range (Hz)**
C_1	20 – 2000
C_2	20 – 1000
C_3	1000 – 2000
C_4	20 – 500
C_5	500 – 1000
C_6	1000 – 1500
C_7	1500 – 2000

Table 3.1: Frequency ranges associated with the pre-trained control filters.

2. **Online noise control**

   During real-time operation, the characteristics of the incoming disturbance are analyzed continuously in order to identify the most appropriate pre-trained control filter. In the proposed implementation, a one-dimensional convolutional neural network (1D CNN), detailed in Section 3.1.2.2, is used to classify the disturbance according to its spectral content and select the corresponding control filter index from the database. Specifically, the CNN is trained, using the synthetic dataset described in Section 3.1.2.2, to associate each one-second frame of incoming noise with the index $i \in \{1, \ldots, 7\}$ of the pre-trained filter C_i that achieves the highest noise reduction for that frame, so that filter selection reduces at runtime to a single forward pass through the network followed by an argmax operation over the output layer.

   Once selected, the filter coefficients are further refined online using the NLMS adaptation mechanism. This hybrid strategy combines the fast response capability of fixed-filter approaches with the adaptability of conventional adaptive filtering methods.

3.1.2.2. CNN-Based SFANC-NLMS Framework

In the proposed hybrid ANC framework, deep learning is employed to perform feature extraction and select the most suitable pre-trained control filter according to the characteristics of the disturbance signal. A one-dimensional convolutional neural network (1D CNN), running on a co-processor, determines the index of the optimal control filter from the SFANC database. The selected filter is then refined online using the NLMS adaptive algorithm, enabling both rapid initialization and adaptive tracking capabilities.

The adopted architecture is derived from the 1D CNN proposed in [43], with modifications introduced to reduce the number of trainable parameters and improve computational efficiency. In particular, the dimensions of the convolutional layers were reduced and the Softmax layer was removed. The resulting architecture is summarized in Table 3.2.

The network begins with a convolutional layer using a kernel size of 5, transforming the input signal into 10 feature maps. A second convolutional layer with a kernel size of 32 further extracts higher-level temporal features while increasing the number of feature maps to 20. A max-pooling layer with a kernel size and stride of 1024 is subsequently applied to reduce feature dimensionality while preserving the dominant information. The extracted features are then flattened and normalized through batch normalization before being processed by two fully connected layers separated by a ReLU activation function. The final output layer produces the index corresponding to the selected control filter.

Layer	**Type**	**Configuration**
1	Convolution	Kernel: 5, Channels: $1 \rightarrow 10$
2	Convolution	Kernel: 32, Channels: $10 \rightarrow 20$
3	Max Pooling	Kernel: 1024, Stride: 1024
4	Flatten	–
5	Batch Normalization	Features: 300
6	Fully Connected	Input: 300, Output: 15
7	ReLU Activation	–
8	Fully Connected	Input: 15, Output: N

Table 3.2: Architecture of the proposed 1D CNN model.

To train the CNN model, a synthetic dataset composed of broadband random noise signals was generated. Each signal has a duration of 1 second and contains randomly varying frequency bands covering the operating range of the ANC system. The corresponding label assigned to each disturbance signal is the index of the control filter achieving the highest noise reduction.

The complete dataset contains 84,000 samples, divided into 80,000 training samples, 2,000 validation samples, and 2,000 testing samples. Model training was performed using the Adam optimizer with a learning rate of 10^{-4}, a batch size of 250, and a total of 30 epochs. To reduce overfitting and maintain a lightweight architecture suitable for embedded implementation, all trainable parameters were regularized using L2 regularization with a factor of 10^{-4}.

The proposed architecture was compared with several one-dimensional neural network models reported in the literature, including the original 1D CNN proposed in [43], the

M6-res architecture presented in [17], and the M3, M5, M11, M18, and M34-res architectures introduced in [44]. Table 3.3 summarizes the classification accuracy and number of trainable parameters for each model.

Network	Test Accuracy	Parameters
Proposed CNN	99.00%	11,687
1D CNN [43]	99.00%	17,147
M6-res [17]	99.40%	210,183
M3 [44]	99.05%	220,423
M5 [44]	99.30%	557,575
M11 [44]	98.95%	1,784,903
M18 [44]	98.50%	3,682,247
M34-res [44]	98.40%	3,982,999

Table 3.3: Comparison between the proposed CNN architecture and existing one-dimensional neural network models.

The results demonstrate that the proposed architecture achieves a competitive classification accuracy while significantly reducing the number of trainable parameters compared to deeper residual architectures. This lightweight configuration makes the proposed CNN particularly suitable for real-time ANC applications implemented on embedded hardware platforms.

3.1.2.3. Hybrid SFANC-NLMS Approach

The system processes incoming noise by sampling it at 16 kHz, capturing one second of audio per frame. A dedicated co-processor analyzes the data using a one-dimensional convolutional neural network (CNN) to determine the most effective noise suppression filter. The selected filter's index is then relayed to a real-time controller, which applies adjustments if necessary. Meanwhile, the controller continuously refines the filter's parameters in response to the evolving sound environment. This setup ensures that the co-processor operates at the frame level while the controller handles real-time signal adjustments. By leveraging advanced adaptive filtering techniques, the system dynamically responds to changing noise conditions, progressively enhancing its noise reduction performance.

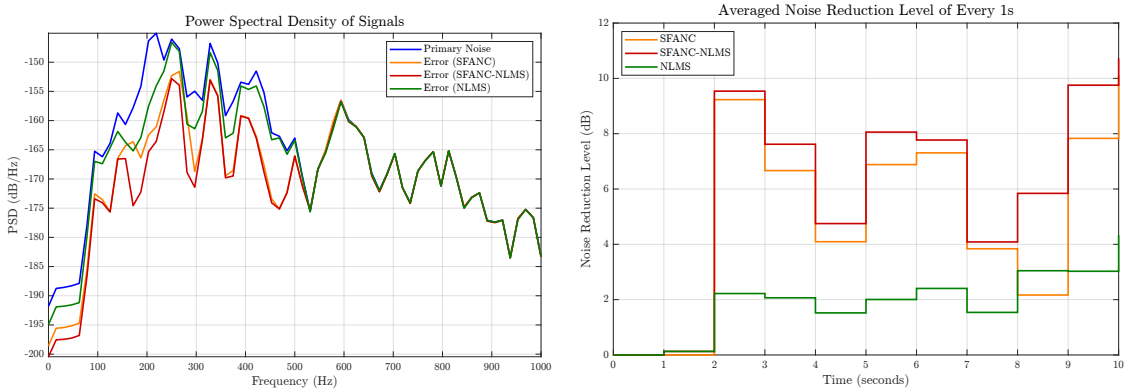
From a stability perspective, the CNN operates only at the frame level to select the most appropriate pre-trained control filter and does not directly perform recursive adaptation of the control coefficients. The online adaptation is instead performed by the NLMS algorithm, whose stability is governed by the choice of the adaptation step size, as established in the adaptive filtering literature. In the experiments conducted in this study, the adap-

tation remained stable for all the considered operating conditions, with no divergence or unbounded growth of the residual error being observed. A formal stability analysis of the complete hybrid CNN–NLMS architecture is beyond the scope of this work.

3.1.3. Results

To evaluate the performance of the proposed algorithm in realistic conditions, experiments were conducted using non-stationary noise sources, including a handheld drill, helicopter, motorbike, and shopping mall recordings. The motorbike noise used in this study is taken from the SONYC dataset [45]. Results are presented in both the frequency domain and in terms of average noise reduction.

The frequency-domain analysis is illustrated in Figure 3.5, which reports the power spectral density (PSD) of the primary and controlled signals, as well as the average noise reduction per second.



(a) PSD comparison (b) Average noise reduction

Figure 3.5: Frequency-domain comparison of NLMS, SFANC, and SFANC-NLMS for motorbike noise.

From the frequency-domain results, the proposed SFANC-NLMS algorithm consistently achieves the lowest spectral energy across most frequency bands. The improvement is particularly significant in the low-frequency range (20–500 Hz), which is the most critical region for active noise control in enclosed acoustic environments. In contrast, both NLMS and SFANC exhibit weaker attenuation and less stable performance across the spectrum.

Quantitatively, Table 3.4 summarizes the noise reduction achieved across all tested noise types. The proposed SFANC-NLMS consistently outperforms both NLMS and SFANC, confirming its improved adaptability to non-stationary disturbances.

Noise type	NLMS	SFANC	SFANC-NLMS
Motorbike	2.70	6.80	**7.73**
Handheld drill	2.70	4.06	**5.43**
Helicopter	2.20	5.03	**6.15**
Shopping mall	3.05	3.81	**5.28**

Table 3.4: Noise reduction performance for different noise types (dB).

3.2. Subjective perception of Noise Reduction Achieved Using a Hybrid Deep Learning-Based Active Noise Control Algorithm

While active noise control (ANC) algorithms are commonly evaluated using objective metrics such as noise reduction in decibels, these indicators do not always reflect human perception of acoustic comfort. In enclosed environments such as vehicle interiors, residual noise characteristics can strongly influence perceived annoyance, even when significant attenuation is achieved.

To address this limitation, this study investigates the subjective perception of noise reduction obtained with ANC algorithms in an encapsulated structure representative of a vehicle interior. In particular, the perceptual effects of a conventional NLMS algorithm and the proposed SFANC-NLMS approach are compared under different noise types and sound pressure levels.

The following research questions (RQ) are addressed:

- **RQ1.** To what extent do different ANC algorithms (NLMS and SFANC-NLMS) influence perceived annoyance across various noise types in a simulated vehicle environment?
- **RQ2.** Does sound level (55, 65, and 72 dB(A)) moderate the relationship between ANC algorithm type and perceived annoyance?
- **RQ3.** How do different noise sources interact with ANC algorithms and sound levels to shape perceived annoyance and acoustic comfort?

3.2.1. Experimental Procedure

Following the acquisition of the experimental signals in the noise box, audio stimuli were generated for the subjective evaluation. Different environmental noises, including mo-

torcycle, street, and train sounds, were reproduced outside the cavity using the external loudspeaker. The ANC algorithms were then applied in real time through the internal canceling loudspeaker. For each noise type, recordings were obtained under three conditions: without ANC, with the NLMS algorithm, and with the proposed SFANC-NLMS algorithm.

The resulting audio signals were calibrated in an Audio Lab using a Head and Torso Simulator (HATS) HEAD acoustics HMS II.3 LN HEC equipped with Sennheiser HD 650 headphones. Prior to stimulus calibration, the HATS was calibrated using a Larson Davis CAL250 sound level calibrator emitting a 251.2 Hz tone at 114 dB. Each stimulus, played from a computer, was individually calibrated by adjusting the playback volume to achieve target levels of 55 dB(A), 65 dB(A), and 72 dB(A). These calibrations were performed first without active noise control, and then adjusted accordingly for the controlled conditions.



(a) (b)

Figure 3.6: Images of the experimental set-up from inside (a) and outside (b) [46].

Calibration was carried out using ArtemiS SUITE 12 software, together with a labO2-V1 playback equalizer and an HDA IV headphone amplifier. To account for variability due to headphone repositioning, each stimulus was calibrated three times, with the headphones removed and repositioned on the HATS between measurements to estimate uncertainty.

Participants first provided demographic information (age and gender) and completed a pre-test assessment including: (1) Individual Noise Sensitivity (INS) using the Weinstein Sensitivity to Sound Scale (WSSN-5) [47], (2) baseline noise annoyance (BNA) based on ISO/TS 15666 [48], and (3) the Perceived Stress Scale (PSS-4) [49]. These measures were used to account for potential confounding factors such as stress and mood in sound

perception [50, 51].

A short training session was then conducted to familiarize participants with the experimental stimuli and the evaluation procedure. Participants were exposed sequentially to the loudest stimulus, the quietest stimulus, and a training signal (train noise at 65 dB(A)). After training, they completed a practice evaluation. All stimuli were generated using FL Studio and presented in a randomized playback sequence.

To assess reliability, an unprocessed 65 dB(A) train noise was used as both the first and last stimulus for each participant. The remaining 27 stimulus conditions were presented in random order. Participants could replay stimuli freely with assistance from the experimenter if required.

Each stimulus was evaluated according to three perceptual dimensions: (1) perceived annoyance (PAY), (2) perceived affective quality (PAQ), and (3) perceived loudness (PLN).

PAY was assessed using ISO 15666, asking participants: *"Thinking about the noise you just heard, how much does it bother, disturb, or annoy you?"* Responses were given on a 5-point scale ranging from "Not at all" to "Extremely".

PAQ was evaluated using the ISO/TS 12913-2 circumplex model [52], which includes eight affective descriptors (eventful, vibrant, pleasant, calm, uneventful, monotonous, annoying, chaotic). Participants indicated their perception by placing a marker on a 2D affective map representing these dimensions.

PLN was assessed using magnitude estimation, where participants rated perceived loudness on a continuous scale from 0 to 100 in response to the question: *"How loud would you say the sound environment is?"* This method allows a fine-grained evaluation of perceived loudness [53].

3.2.2. Data Analysis

The collected data were processed to compute mean scores and standard deviations for INS, PSS-4, and BNA. For INS and PSS-4, individual scores were obtained by summing all questionnaire responses and dividing by the number of items (5 for INS and 4 for PSS-4). Group-level means were then computed across participants, with standard deviations used to quantify inter-individual variability.

For BNA, scores were computed separately for each noise type (e.g., traffic, airplane, train) using the same averaging procedure. Group-level statistics were then calculated per noise category to characterize perceived baseline annoyance across different acoustic

conditions.

Statistical analyses were conducted using a three-way repeated measures ANOVA (3WR-ANOVA) to assess the effects of noise type, noise level, and ANC condition on perceived annoyance (PAY) and perceived loudness (PLN). Post-hoc comparisons were performed using Tukey's HSD test.

Prior to ANOVA, assumptions of normality, homogeneity of variance, and sphericity were verified using the Shapiro–Wilk test [54], Levene's test [55], and Mauchly's test [56], respectively. All assumptions were satisfied.

For PAQ, a three-way repeated measures permutational multivariate ANOVA (3WR-PERMANOVA) was used due to violations of multivariate normality (confirmed by Mardia's test [57]). The analysis considered the PAQ coordinates (ISOPL and ISOEV) as dependent variables, with noise type, noise level, and ANC condition as factors. Post-hoc pairwise comparisons were conducted using the pairwise.adonis() function to identify significant group differences across factor combinations.

3.2.3. Discussion

This section discusses the findings with respect to the research questions introduced. The impact of the ANC algorithms on perceived annoyance (RQ1) is first addressed, followed by the role of sound level under different ANC conditions (RQ2), and finally the combined effects of noise type, ANC strategy, and sound level (RQ3).

3.2.3.1. Effect of ANC algorithms on perceived annoyance

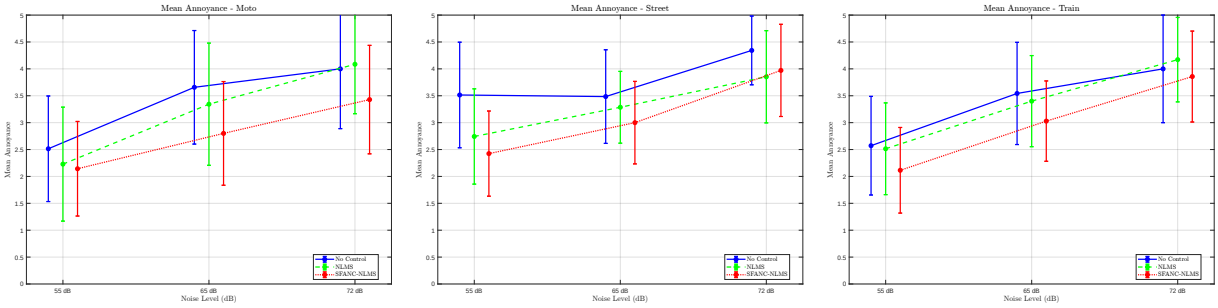
The results indicate that the proposed SFANC-NLMS algorithm consistently reduces perceived annoyance compared to the conventional NLMS approach, addressing RQ1. This improvement is observed across all tested noise types and sound levels, with particularly strong effects for motorbike and train noise, where SFANC-NLMS achieves the lowest annoyance ratings.

Overall, SFANC-NLMS demonstrates greater robustness to complex and non-stationary acoustic environments. While NLMS provides moderate improvement, its performance is less stable across conditions. In contrast, the hybrid structure of SFANC-NLMS, combining pre-trained filter selection with adaptive refinement, results in more consistent perceptual benefits.

Noise level also has a clear influence on perceived annoyance, with higher sound pressure levels (e.g., 72 dB(A)) systematically increasing annoyance regardless of the ANC condi-

tion. However, SFANC-NLMS maintains a relative advantage across all levels, suggesting that its benefit is not limited to low-noise scenarios.

These trends are illustrated in Figure 3.7, where SFANC-NLMS generally yields the lowest perceived annoyance scores across all noise types and levels.



Motorbike Street Train

Figure 3.7: Mean perceived annoyance (PAY) and standard deviations for different ANC conditions across noise levels.

3.2.3.2. Effect of sound level on perceived annoyance under different ANC conditions

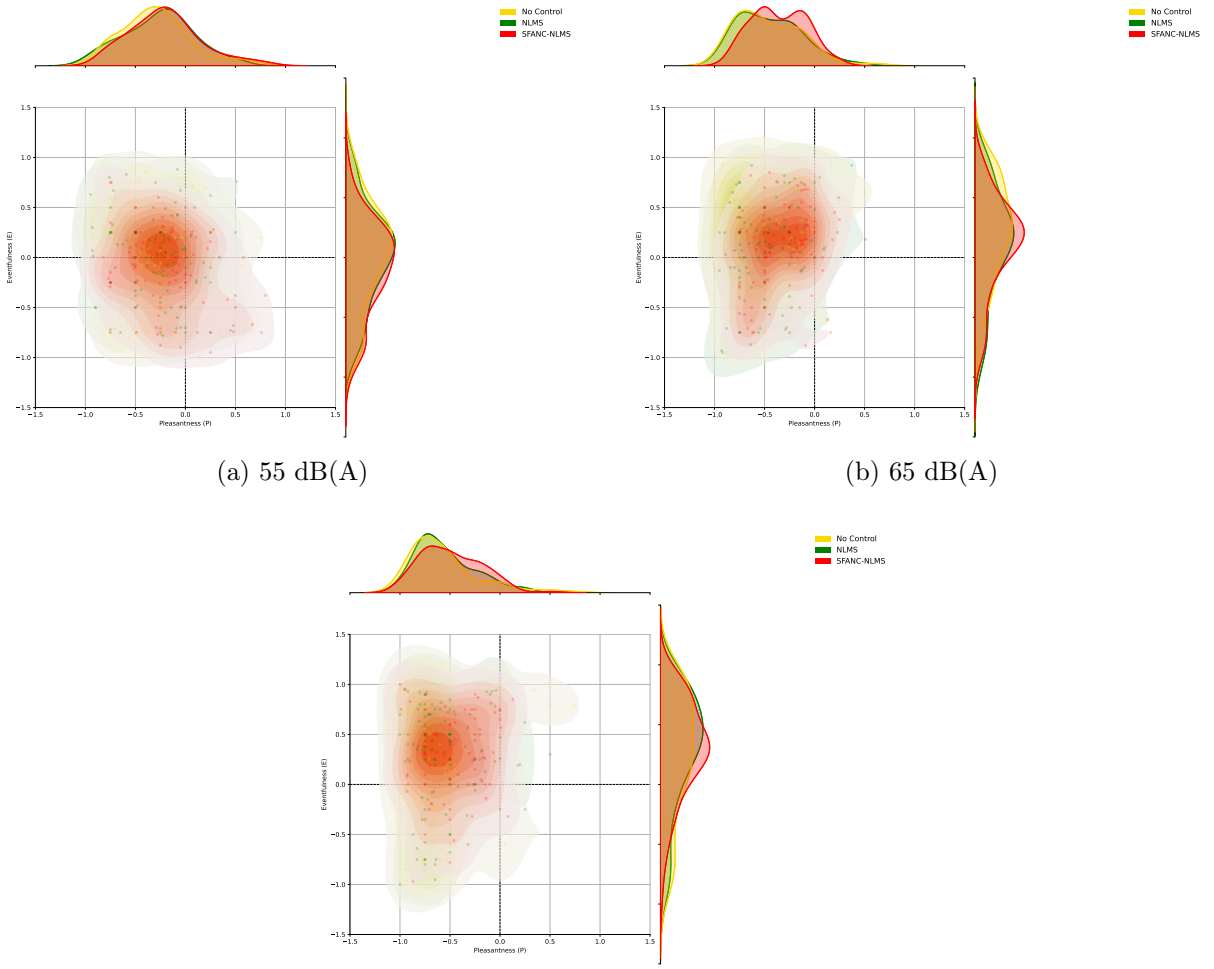
The results show that sound level (55, 65, and 72 dB(A)) plays a key role in shaping perceived annoyance, addressing RQ2. In general, higher sound levels lead to higher annoyance ratings, even when ANC is applied. This confirms that sound intensity remains a dominant factor in perceived discomfort.

The effect of sound level is most pronounced at 72 dB(A), where differences between ANC conditions become more evident. However, while sound level amplifies perceived annoyance, it does not fundamentally change the relative behavior of the ANC algorithms, which remain consistent across conditions.

These trends are further supported by the PAQ analysis shown in Figure 3.8. The circumplex representation indicates a systematic shift in perceptual quality as sound level increases.

As sound level increases, responses shift toward the upper-left region of the circumplex space, corresponding to more chaotic and less pleasant perceptions. This trend is consistent with the intrinsic relationship between sound intensity and perceived annoyance.

Regarding the ANC condition, SFANC-NLMS consistently shifts the distribution toward the lower-right region of the circumplex map, indicating a more pleasant and calmer



(a) 55 dB(A)

(b) 65 dB(A)

(c) 72 dB(A)

Figure 3.8: Circumplex representation of perceived affective quality (PAQ) across sound levels. ISOPL (pleasantness) is plotted against ISOEV (eventfulness) for 55, 65, and 72 dB(A).

soundscape. This suggests that, beyond reducing annoyance, the proposed method also improves the overall affective quality of the acoustic environment.

3.2.3.3. Interaction between noise sources, ANC condition, and sound levels

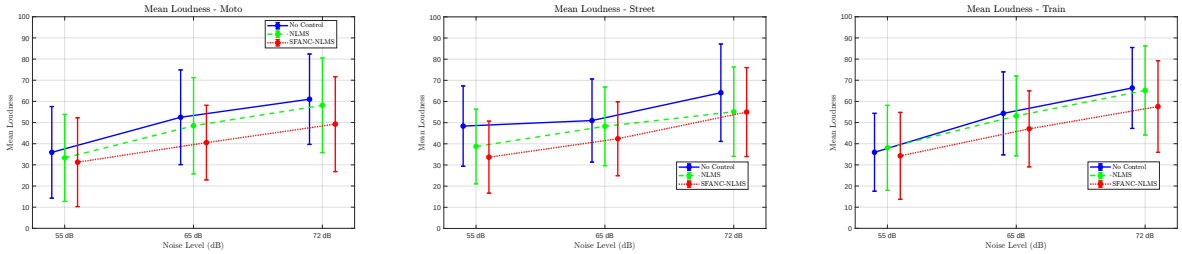
The results highlight clear interactions between noise sources, ANC condition, and sound levels, addressing RQ3. Overall, the influence of ANC on perceived annoyance depends on the type of noise and its spectral characteristics, with additional modulation from sound level.

In particular, street and train noises exhibit a stronger sensitivity to the ANC condi-

tion, where SFANC-NLMS consistently achieves lower perceived annoyance compared to NLMS. For motorbike noise, SFANC-NLMS also provides improved performance; however, the interaction between noise type and ANC algorithm is less pronounced, suggesting that highly non-stationary or broadband disturbances are less dependent on algorithmic differences.

Across all noise types, higher sound levels systematically increase perceived annoyance, confirming the dominant role of sound intensity. Nevertheless, SFANC-NLMS maintains its advantage across all tested conditions, indicating robust performance in heterogeneous acoustic environments.

These findings are further supported by the perceived loudness (PLN) results shown in Figure 3.9. SFANC-NLMS consistently yields the lowest loudness ratings across all noise types and sound levels, reinforcing its ability to improve overall acoustic comfort.



Motorbike Street Train

Figure 3.9: Mean perceived loudness (PLN) and standard deviations across noise levels for different ANC conditions and noise types.

3.3. Hybrid Generative Deep Learning-Based ANC for Encapsulated Structures with Openings under Nonlinearities

Encapsulated structures with openings are widely used in practical applications where thermal regulation and airflow are required. However, these openings significantly degrade passive noise insulation and make effective acoustic control more challenging, especially at low frequencies where passive solutions are inherently limited.

Although active noise control (ANC) provides a promising alternative, its implementation in such configurations remains difficult due to geometric constraints and the reduced effectiveness of conventional actuator arrangements. In addition, real-world ANC systems are affected by nonlinearities introduced by loudspeakers and other electroacoustic

components, which further limit the performance of classical adaptive algorithms.

These challenges motivate the development of a hybrid ANC approach capable of handling both structural openness and system nonlinearities. In this section, the proposed GFANC-THFxNLMS algorithm, which leverages deep learning to mitigate the nonlinearities introduced by loudspeakers in ANC applications for encapsulated structures with openings, is presented and detailed.

3.3.1. Environment of control

The identification experiments were conducted in a laboratory room measuring 5.8 m in length and 3.5 m in width. Although the walls are partially covered with sound-absorbing materials, the presence of various equipment and objects results in acoustic conditions that are closer to a realistic operational environment than those of an anechoic chamber.

The experimental setup consists of a rectangular steel plate mounted on a concrete enclosure that houses the primary noise source, as shown in Figure 3.10. A Behringer EUROLIVE B208D loudspeaker is used as the primary excitation source, generating band-limited random noise with frequencies up to 1500 Hz and a maximum sound pressure level (SPL) of 113 dB. The secondary source is a JBL Stage2 424 loudspeaker located inside the enclosure and positioned centrally behind the opening. The control hardware is implemented using the Analog Devices SC598 with SHARC+ core and the EVSOMCRR-EZKIT development board [58].

The steel plate has a density of $\rho = 7850\,\mathrm{kg/m^3}$ and a Young's modulus of $E = 210 \times 10^9\,\mathrm{Pa}$. Its dimensions are $a = 0.42\,\mathrm{m}$ (length), $b = 0.39\,\mathrm{m}$ (width), and $h = 0.001\,\mathrm{m}$ (thickness). This configuration is representative of encapsulated structures with openings commonly encountered in engineering applications.

The primary and secondary acoustic paths are identified using the Normalized Least Mean Square (NLMS) algorithm. The signals are recorded at a sampling rate of 48 kHz and then downsampled to 3 kHz to reduce computational complexity, since the frequency content of interest lies below this limit. A filter length of 300 coefficients is used for both paths.

The identified responses reveal distinct characteristics for each path. The primary path exhibits strong resonant behavior, particularly below 350 Hz, due to the coupled dynamics between the vibrating plate and acoustic leakage through the opening [59]. In contrast, the secondary path, driven by the internal loudspeaker, shows a smoother and more attenuated response, indicating a more direct acoustic propagation with limited structural coupling. Both paths display stable and physically consistent behavior, confirming the

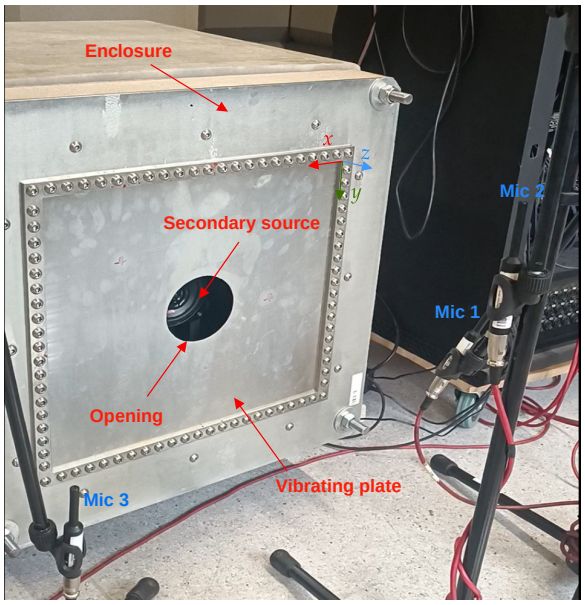


Figure 3.10: Overview of the experimental rig showing a vibrating steel plate with a central circular opening of radius 0.05 m. Microphones (Mic 1, Mic 2, and Mic 3) are positioned at different locations within the enclosure. Mic 1 is used as the error microphone, while Mic 2 and Mic 3 serve as observation points. The primary noise source is placed outside the enclosure, and the secondary loudspeaker is positioned behind the opening.

validity of the identified models.

3.3.2. The GFANC-THFxNLMS algorithm

This section presents the proposed GFANC-THFxNLMS algorithm, which integrates a deep learning-based filter weighting mechanism with a nonlinear adaptive control stage. The method is designed to operate effectively in the encapsulated structure described in the previous section. The overall architecture is illustrated in Figure 3.11.

3.3.2.1. THFxNLMS adaptive control stage

The THFxNLMS algorithm extends the standard FxNLMS approach by introducing a nonlinear error transformation to improve robustness against system nonlinearities [61]. Let $x(n) = [x(n), x(n-1), \ldots, x(n-L+1)]^T$ be the reference signal vector and $w(n)$ the adaptive filter coefficients. The control output is:

$$y(n) = w^T(n)x(n). \tag{3.1}$$

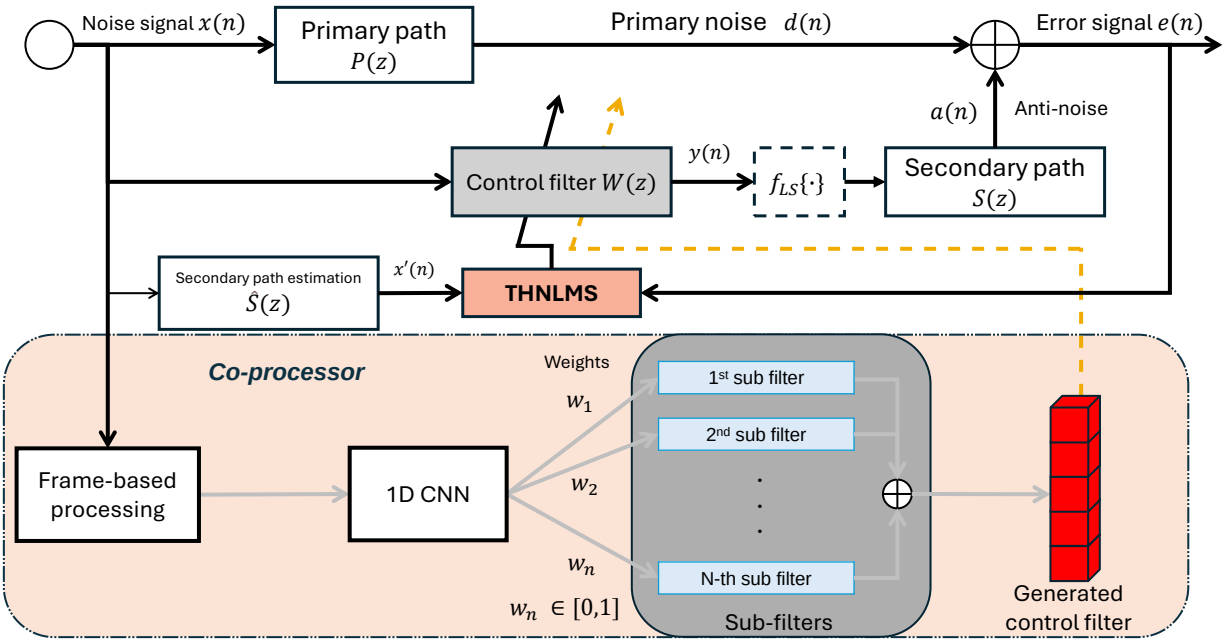


Figure 3.11: Block diagram of the hybrid GFANC-THFxNLMS algorithm. f_{LS} models the loudspeaker nonlinearity. A 1D CNN running on a co-processor selects the most suitable pre-trained control filter, which is then refined online using the THFxNLMS algorithm. The arrow from the "Generated control filter" block to $W(z)$ denotes a one-time initialization: at the start of each processed signal, $W(z)$ is initialized with the CNN-generated filter, after which it is continuously updated sample-by-sample by the THFxNLMS adaptation rule, without further frame-wise resets from the CNN [60].

The residual error is given by:

$$e(n) = d(n) - s(n) * y(n), \tag{3.2}$$

where $d(n)$ is the primary disturbance and $s(n)$ is the secondary path impulse response. To improve robustness, the error is processed through a hyperbolic tangent function:

$$f(e(n)) = \tanh(b(n)e(n)), \tag{3.3}$$

where $b(n)e(n)$ denotes the product of the time-varying compression parameter $b(n)$ and the instantaneous error $e(n)$, and not a separate variable. This bounded transformation reduces the effect of large error peaks and improves stability in nonlinear acoustic

conditions. The resulting update rule of THFxNLMS is:

$$w(n+1) = w(n) + \mu \frac{\tanh(b(n)e(n)) \left[1 - \tanh^2(b(n)e(n))\right]}{2\|x'(n)\|^2 + \epsilon} x'(n), \tag{3.4}$$

where $x'(n)$ is the filtered reference signal, μ is the step size, and ϵ is a small regularization constant.

The term $\left[1 - \tanh^2(b(n)e(n))\right]$, which is the derivative of the tanh transformation, governs the adaptability of the algorithm: it remains close to 1 when $|b(n)e(n)|$ is small, so the update behaves like standard normalized LMS and stays fully sensitive to small errors, but it decays toward 0 as $|b(n)e(n)|$ grows large, automatically shrinking the effective step size whenever the error is large. This built-in, gradient-clipping-like behavior suppresses the influence of large error transients, such as those caused by loudspeaker nonlinearities, without requiring an explicit outer clipping rule.

The compression parameter b controls the nonlinearity strength of the transformation, and therefore the location of the transition between these two regimes. A large b shifts this transition toward smaller error magnitudes, improving outlier rejection but slowing convergence, while a small b keeps a wider near-linear, sensitive regime, accelerating adaptation but reducing robustness.

To improve adaptability, b is updated online as:

$$b(n) = \frac{\rho}{1 + \exp(-P_e(n))}, \tag{3.5}$$

with

$$P_e(n) = \lambda P_e(n-1) + (1-\lambda)e^2(n), \tag{3.6}$$

where $\rho \in [0, 1]$ and $\lambda \in (0.9, 1)$. Since $P_e(n)$ is an exponentially-smoothed estimate of the instantaneous error power, $b(n)$ automatically increases when the error power rises, strengthening outlier rejection, and decreases when the error power is low, restoring sensitivity for fine adaptation.

This adaptive mechanism allows the algorithm to adjust its nonlinear behavior according to the instantaneous noise level, improving overall stability and convergence in non-stationary and nonlinear environments.

Regarding the stability of the overall adaptation, a sufficient condition on the step size μ

can be derived from the weight update rule above:

$$0 < \mu < \frac{1}{\lambda_{\max} D_s}, \tag{3.7}$$

where D_s is the group delay of the secondary path, estimated from its impulse response $s(n)$, and $\lambda_{\max}$ is the maximum eigenvalue of the autocorrelation matrix of the filtered reference signal $x'(n)$, i.e., $\lambda_{\max} = \max \operatorname{eig}\big(E[x'(n)x'^T(n)]\big)$. This inequality provides a verifiable upper bound ensuring that the adaptive update remains stable under the nonlinear transformation; in practice, μ is chosen conservatively below this bound based on the expected signal and secondary-path characteristics.

3.3.2.2. GFANC Method

The Generative Fixed-Filter Active Noise Control (GFANC) method introduces a noise reduction strategy based on a pre-trained set of broadband control filters. Starting from an initially optimized broadband filter, a database of sub-control filters is constructed, enabling a weighting mechanism that dynamically generates an optimal control filter during operation [18]. The method operates in two stages: (1) offline filter training and decomposition, and (2) real-time noise control.

During the first stage, the ANC system is used to cancel a broadband primary noise $d(n)$ containing the relevant frequency components. The optimal control filter, obtained using the THFxNLMS algorithm, is used as the pre-trained broadband filter. Once obtained, this filter is decomposed into sub-control filters using a perfect reconstruction filter bank approach [62].

Given a pre-trained control filter c with L coefficients, its frequency-domain representation is computed using the discrete Fourier transform (DFT):

$$G = F_L \, c$$

where F_L is the DFT matrix. Using the conjugate symmetry property of real-valued signals, the spectrum is decomposed into C sub-bands:

$$G = \sum_{m=1}^{C} G_m$$

Each sub-filter is defined by isolating frequency components within each band, ensuring

perfect reconstruction. The time-domain sub-filters are obtained via inverse DFT:

$$c_m = F_L^{-1} G_m$$

The final control signal is reconstructed as:

$$y(n) = \sum_{m=1}^{C} x^T(n) c_m$$

This decomposition preserves the optimal characteristics of the original filter while enabling flexible control through a reduced set of fixed sub-filters [18].

Concerning the 1D CNN, motivated by [16], the proposed 1D Convolutional Neural Network (CNN) is designed to process 1-second time-domain signals in order to extract relevant spectral-temporal features and generate adaptive control filter weights for the GFANC framework.

The input consists of a one-dimensional waveform, which is processed through successive convolutional and fully connected layers. The first convolutional layer applies 10 filters with a kernel size of 5, capturing local temporal dependencies in the signal. This is followed by a second convolutional layer with 20 filters and a kernel size of 32, enabling the extraction of higher-level representations.

A Max-Pooling operation (kernel size 1024, stride 1024) is then applied to reduce temporal dimensionality while preserving the most informative components. The resulting feature maps are flattened into a 40-dimensional vector and normalized using batch normalization to improve training stability.

The extracted features are then passed through two fully connected layers. The first layer maps the feature vector to a 4-dimensional space followed by a Tanh activation function, introducing non-linearity. The final layer outputs 4 continuous values in the range $[0, 1]$, which correspond to the adaptive weighting coefficients used to combine the sub-control filters in the GFANC framework.

Overall, the network learns a frequency-to-gain mapping rather than directly modeling the noise signal. This allows the CNN to infer optimal filter combination weights based on the spectral characteristics of the reference input.

The detailed architecture of the network is summarized in Table 3.5.

The model is trained as a regression network using a Mean Squared Error (MSE) loss function. The dataset consists of 80,000 training samples and 2,000 validation samples,

Table 3.5: Architecture of the 1D CNN Model.

Layer	Type	Parameters	Output Shape
Conv1	Conv1D	In: 1, Out: 10, Kernel: 5, Stride: 1	(10, 15996)
Conv2	Conv1D	In: 10, Out: 20, Kernel: 32, Stride: 1	(20, 15965)
MaxPool	MaxPool1D	Kernel: 1024, Stride: 1024	(20, 15)
Flatten	Flatten	-	(40)
BatchNorm	BatchNorm1D	Num features: 40	(40)
Linear1	Fully Connected	In: 40, Out: 4	(4)
Tanh	Activation	-	(4)
Linear2	Fully Connected	In: 4, Out: 4	(4)

each corresponding to 1-second signals filtered within randomly selected frequency bands between 20 Hz and 1480 Hz. The target outputs are four continuous-valued coefficients in the range $[0, 1]$, representing the optimal sub-filter combination weights.

Optimization is performed using the Adam optimizer with a learning rate of 2×10^{-4} and a weight decay of 10^{-4}. A learning rate scheduler reduces the learning rate by a factor of 0.5 every 5 epochs to ensure stable convergence. Xavier initialization is used for convolutional layers to improve training stability.

The final trained model contains 6.74K parameters and requires 19.33 million MACs per 1-second frame (3 kHz sampling rate), as summarized in Table 3.6. This relatively low computational cost allows real-time deployment on embedded architectures such as the ADSP-SC598, where the model can be fully stored in on-chip memory, minimizing latency and memory transfers.

Table 3.6: Computational complexity of the proposed 1D CNN model.

Component	Parameters	MACs per frame
1D CNN	6.74 K	19.33 M

3.3.2.3. GFANC-THFxNLMS Approach

The system processes incoming noise by sampling it at 3 kHz, capturing one second of audio per frame. As illustrated in Fig. 3.11, a dedicated coprocessor analyzes each frame using the GFANC framework. In this configuration, a one-dimensional convolutional neural network (1D CNN) is employed to generate a set of weights. These CNN-predicted weights are used to construct an initial optimal control filter that is tailored to the spectral characteristics of the current noise frame. Once the control filter is generated, it is

transferred to the real-time controller, where the THFxNLMS algorithm takes over. Importantly, is to fine-tune the coefficients of the generated control filter during continuous operation. THFxNLMS is specifically designed to handle system nonlinearities through its nonlinear error formulation and adaptive gradient scaling, allowing the controller to compensate for secondary-path distortions and nonlinear loudspeaker–acoustic effects that cannot be captured at the frame level. This separation of roles is central to the proposed architecture: the CNN operates at the frame level to provide a robust and noise dependent initialization of the control filter, while THFxNLMS performs sample level adaptation to explicitly address nonlinearities and time-varying dynamics in the ANC system. By combining the generative fixed filter approach with the THFxNLMS algorithm, the system ensures dynamic adaptation, progressively improving the noise reduction performance in real time. Unlike the algorithm proposed in [17], once the initial control filter is generated, THFxNLMS continues to adapt its coefficients throughout the remainder of the signal without requiring frame-wise CNN updates. Table 3.7 summarizes the step-by-step procedure of the proposed algorithm, where $\mathbf{H}$ denotes the subfilters dataset and $\hat{\mathbf{w}}_{\text{CNN}}$ represents the weights predicted by the 1D CNN.

Table 3.7: Step-by-step description of the GFANC-THFxNLMS algorithm with corresponding equations.

Step Description	**Corresponding Equation**
1D CNN predicts weights	$\hat{\mathbf{w}}_{\text{CNN}} = \text{CNN}(x(n))$
Generated control filter	$\mathbf{w} = \hat{\mathbf{w}}_{\text{CNN}} \times \mathbf{H}$
Filtered reference signal	$\mathbf{x}'(n) = \hat{s}(n) * x(n)$
Compute controller output	$y(n) = \mathbf{w}^T(n)\mathbf{x}(n)$
Compute error signal	$e(n) = d(n) - s(n) * f_{SEF}(y(n))$
Update error power estimate	$P_e(n) = \lambda P_e(n-1) + (1-\lambda)e^2(n)$
Compute gradient scaling factor	$b(n) = \frac{\rho}{1+e^{-P_e(n)}}$
Compute weighted loss	$\text{loss}(n) = \dfrac{\tanh^2\big(b(n)e(n)\big)}{2 \cdot \text{power}(n) + \varepsilon}$
Update filter weights	$\mathbf{w}(n+1) = \mathbf{w}(n) - \mu\psi(e(n))\mathbf{x}'(n)$ where $\psi(e) = \dfrac{\tanh(b(n)e(n))[1-\tanh^2(be)]}{2\|\mathbf{x}'(n)\|^2 + \varepsilon}$

The stability of the proposed GFANC-THFxNLMS approach relies primarily on the THFxNLMS adaptive stage, while the CNN operates at the frame level and provides only the initialization of the control filter. The bounded hyperbolic tangent transformation used in the THFxNLMS update limits the influence of large error values and contributes to the robustness of the adaptation in the presence of nonlinearities. In addition, the

normalization of the filtered reference signal and the regularization term ε prevent excessively large parameter updates. Experimentally, no divergence or unbounded growth of the residual error was observed under the considered operating conditions. A formal theoretical stability analysis of the complete CNN-based THFxNLMS architecture is beyond the scope of this work.

3.3.2.4. Computational complexity

This section analyzes the computational complexity of the proposed GFANC-THFxNLMS algorithm.

The proposed method combines a CNN-based filter generation stage with the adaptive THFxNLMS update mechanism. The CNN inference introduces a fixed computational cost denoted by C_{CNN}, which accounts for the forward pass of the network used to generate the sub-control filter weights.

After the CNN stage, the control filter is constructed through a weighted combination of sub-control filters, requiring CL multiplications and $C(L-1)$ additions/subtractions, where C is the number of subfilters and L is the filter length.

The adaptive THFxNLMS component further contributes $8L+11$ multiplications and $4L+7$ additions/subtractions per iteration. This includes the filtered reference computation, nonlinear error transformation, and weight update process.

Overall, the total computational complexity of the GFANC-THFxNLMS algorithm is given by: $C_{\text{CNN}}+CL+8L+11$ multiplications and $C_{\text{CNN}}+CL+4L+8$ additions/subtractions.

Although the CNN inference and filter combination introduce additional computational cost, the algorithm remains suitable for real-time implementation due to its compact network design and efficient frame-based processing strategy. Memory usage is minimized by storing the CNN parameters on-chip and reusing input buffers across successive frames, thereby reducing latency and memory bandwidth requirements in embedded systems.

3.3.3. Results

This section presents simulation results for a single-channel active noise control (ANC) system. The sampling rate is set to 3 kHz, and the control filter length is 256 taps. The primary and secondary paths are defined using the impulse responses introduced in Section 3.3.1. The system employs four pre-trained sub-control filters, and the THFxNLMS algorithm is used to derive the optimal control filter with a step size of 0.1.

Table 3.8: Complexity analysis of the GFANC-THFxNLMS algorithm.

Operation	**Dimensions**	$*$ **Multiplications**	$+$ / $-$ **Additions/Subtractions**
1D CNN	Input signal	C_{CNN}	C_{CNN}
$\mathbf{w} = \hat{\mathbf{w}}_{\text{CNN}}\mathbf{H}$	$1 \times C, C \times L$	CL	$C(L-1)$
$x_f(n)$	$\mathbb{R}^{L\times 1}$	L	1
$y(n)$	$\mathbb{R}^{L\times 1}$	L	$L-1$
$e(n)$	Scalar	1	1
$P_e(n)$	Scalar	2	1
$b(n)$	Scalar	2	1
loss	Scalar	6	3
$\mathbf{w}(n+1)$	$\mathbb{R}^{L\times 1}$	$4L+2$	$3L+1$
Total		$C_{\text{CNN}} + CL + 8L + 11$	$C_{\text{CNN}} + CL + 4L + 8$

Loudspeaker saturation is a well-known source of nonlinear distortion in ANC systems [61]. This phenomenon has been widely studied in the literature [20, 22] and is commonly modeled using the scaled error function (SEF) [63]:

$$f_{\text{SEF}}(y) = \int_0^y e^{-\frac{2z}{\eta^2}}\, dz, \tag{3.8}$$

where y denotes the loudspeaker input and η controls the level of nonlinearity. This model captures saturation effects typically observed in practical loudspeakers. As $\eta \to \infty$, the system behaves linearly, whereas small values of η correspond to strong saturation effects.

To evaluate robustness under nonlinear distortions, three different loudspeaker conditions are considered:

- $\eta = 0.5$ (severe nonlinearity),
- $\eta = 5$ (moderate nonlinearity),
- $\eta = \infty$ (linear case).

Figure 3.12 illustrates the effect of the nonlinearity parameter η. Large values correspond to nearly linear behavior, where performance is mainly driven by the adaptive filter. For moderate values, realistic saturation effects appear while maintaining controllable behavior. For small values, strong nonlinear distortion significantly affects control performance.

In the evaluation phase, both trained and unseen loudspeaker nonlinearities are considered. It is important to note that the sub-control filters are pre-trained using signals that pass through nonlinear loudspeaker models, ensuring that nonlinear effects are embedded in the learned representations. The proposed 1D CNN then combines these nonlinear-

aware sub-filters, while the THFxNLMS stage further refines the control filter online.

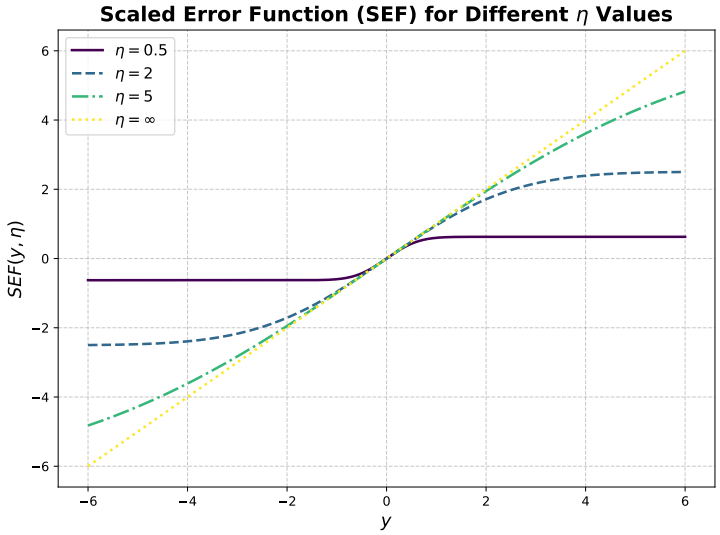


Figure 3.12: Scaled error function (SEF) for different values of η.

To evaluate the effectiveness of the proposed approach, real-world recorded non-stationary noise signals are used. The system is tested under varying noise conditions and nonlinear settings. A time-varying study is also conducted to assess robustness under dynamic acoustic environments with $\eta = 5$.

A 10-second hood fan signal is used, where the signal-to-noise ratio (SNR) of the primary path varies over time. The SNR is kept constant during the first segment, set to 5 dB in the second segment, and increased to 8 dB in the final segment.

The results show that all algorithms achieve noise attenuation, but the proposed GFANC-THFxNLMS consistently produces the lowest error amplitude. This improvement is also confirmed in the frequency domain and in the noise reduction analysis per second.

In the frequency domain, the proposed method achieves stronger attenuation in the low-frequency range (approximately 100–350 Hz). Over time, it is observed that THFxNLMS gradually approaches the performance of GFANC, highlighting the complementary nature of both approaches in terms of convergence and adaptability. The proposed GFANC-THFxNLMS benefits from both fast adaptation and robust filter generation, leading to the best overall performance.

The average noise reduction achieved by each method is 3.63 dB for THFxNLMS, 3.35 dB for GFANC, 3.12 dB for SFANC, and 4.83 dB for GFANC-THFxNLMS.

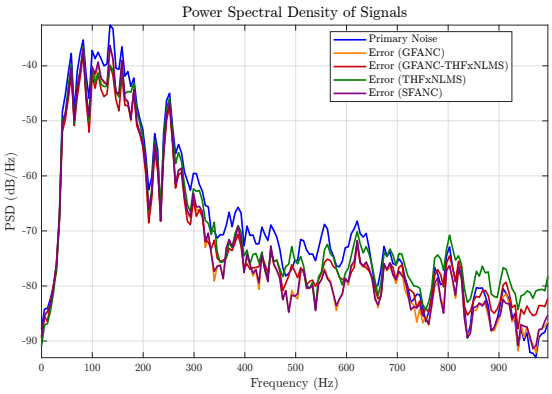


Figure 3.13: PSD comparison

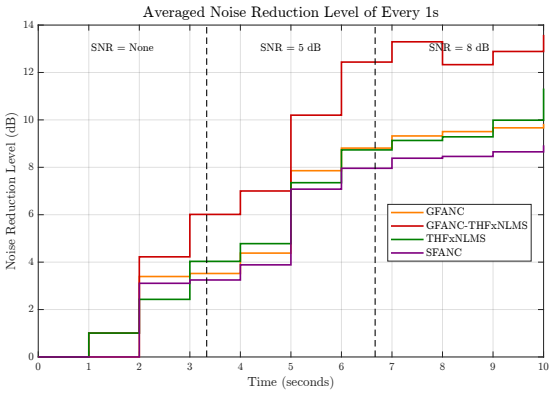


Figure 3.14: Noise reduction per second

Figure 3.15: Comparison of spectral performance and temporal noise reduction under varying SNR conditions.

4 | Deep transfer reinforcement learning based active noise control in vehicle interior

Active noise control inside vehicle cabins remains particularly challenging due to the presence of nonlinear and time-varying acoustic conditions generated by engine operation, road–tire interaction, and structural vibrations. These complex vibro-acoustic behaviors significantly limit the performance of conventional adaptive ANC algorithms, especially under real driving conditions.

Recent advances in deep learning have shown promising capabilities for modeling nonlinear acoustic systems, while reinforcement learning enables autonomous adaptation through interaction with the environment. In this section, a deep transfer reinforcement learning-based active noise control algorithm for vehicle interiors is presented. The proposed approach combines a modified deep neural ANC architecture with a reinforcement learning strategy to improve robustness and adaptability in realistic automotive acoustic environments.

4.1. Environment of control

The acoustic environment used in this study was identified experimentally inside a sedan vehicle equipped with a four-cylinder gasoline engine. Figure 4.1 presents views of the experimental vehicle from both the interior and exterior.

All signals were recorded with a sampling frequency of 3 kHz. The secondary paths were identified experimentally using the Multiple Exponential Sine Sweep Method (MESM) [64] for a five-speaker and four-microphone ANC configuration. The setup included the four door-mounted woofers and the factory-installed subwoofer located in the trunk of the vehicle. Four Mikrotek Geffel M380 microphones were positioned at the seating locations corresponding to the driver seat, front passenger seat, and the two rear passenger seats.



(a) (b)

Figure 4.1: Experimental vehicle used for the measurements: (a) interior view and (b) exterior view.

To account for variations in the acoustic environment, measurements were conducted under three occupancy conditions: empty cabin, one occupant, and two occupants. These configurations were used to capture the influence of passenger-induced acoustic loading on the secondary path dynamics.

The primary noise dataset was recorded during real driving conditions on German roads, including both Autobahn and urban traffic scenarios. Approximately two hours of in-vehicle noise were collected using the same microphone configuration and a PAK MKII acquisition frontend. From this dataset, 10,000 one-second signals were extracted for training. A separate five-minute recording segment was excluded from training and reserved exclusively for evaluation to ensure unbiased testing of the proposed algorithm on unseen data.

4.2. Deep learning based active noise control

4.2.1. Filtered-x Modified DNoiseNet Algorithm

The proposed FxMDNoisenet algorithm is a deep learning-based feedback control strategy designed to directly generate an anti-noise signal from the estimated disturbance.

The block diagram of the FxMDNoisenet algorithm is shown in Figure 4.2. In this representation, $d(n)$ denotes the primary disturbance, $a(n)$ the anti-noise signal generated through the secondary path, $y(n)$ the control signal produced by FxMDNoisenet, $\hat{d}(n)$ the estimated disturbance, and $e(n)$ the residual error. $S(n)$ and $\hat{S}(n)$ represent the true secondary path and its estimate, respectively.

At each discrete-time instant n, the input of MDNoisenet consists of a sequence of L past

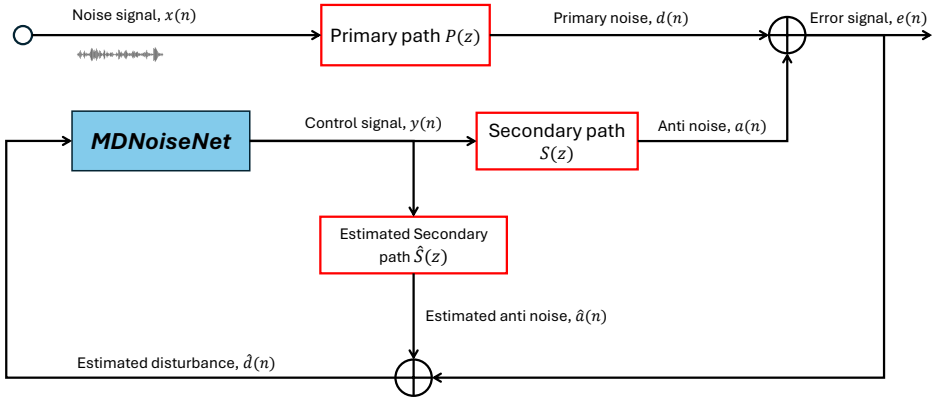


Figure 4.2: FxMDNoisenet for feedback ANC. The network takes the estimated disturbance $\hat{d}(n)$ as input and generates the control signal $y(n)$, which is filtered through the secondary path $S(n)$ to produce the anti-noise $a(n)$. The residual $e(n)$ is obtained by subtracting $a(n)$ from the disturbance $d(n)$.

samples of the estimated disturbance:

$$\hat{\mathbf{d}}_L(n) = \big(\hat{d}(n-L+1), \hat{d}(n-L+2), \ldots, \hat{d}(n)\big)^T, \tag{4.1}$$

and the control signal is generated as

$$y(n) = \text{MDNoisenet}\big(\hat{\mathbf{d}}_L(n)\big). \tag{4.2}$$

The resulting anti-noise signal is obtained after propagation through the secondary path:

$$a(n) = S(n) * y(n), \tag{4.3}$$

and the residual error signal is computed as

$$e(n) = d(n) - a(n), \tag{4.4}$$

where $*$ denotes linear convolution. The estimated disturbance used as network input is recursively updated as

$$\hat{d}(n) = e(n) + \hat{S}(n) * y(n), \tag{4.5}$$

where $\hat{S}(n)$ denotes the estimated secondary path. This recursive formulation ensures that FxMDNoisenet operates under a realistic feedback signal that incorporates both the residual error and the secondary path dynamics, enabling stable adaptation to the time-varying acoustic environment.

4.2.2. MDNoiseNet architecture

Motivated by [22], the proposed Modified DNoiseNet (MDNoiseNet) is designed to process raw time-domain acoustic signals under real-time constraints. It combines causal and atrous scaled convolutions with residual and skip connections to enable efficient multi-scale temporal feature extraction while preserving causality.

The input is a single-channel waveform processed by an initial causal 1D convolution (kernel size 2), followed by a Tanh activation. This layer extracts low-level temporal features and serves as the initial skip connection.

The core of the network consists of stacked Atrous Scaled Convolution (ASC) modules with increasing dilation factors, allowing progressive expansion of the receptive field. Each module includes SeLU activations, 1×1 pointwise convolutions, and residual/skip connections to improve gradient flow and aggregate multi-scale information.

After convolutional processing, skip features are regularized using dropout (10%) and reshaped into a compact representation, which is then processed by a two-layer RNN (hidden size 2) to capture remaining temporal dependencies. The final hidden state is passed through two fully connected layers, producing a single scalar output corresponding to the control signal.

Overall, MDNoiseNet integrates causal convolutions, multi-scale atrous filtering, and recurrent modeling to effectively capture temporal dynamics in acoustic signals, making it suitable for real-time active noise control applications.

The detailed architecture of the proposed network is illustrated in Figure 4.3.

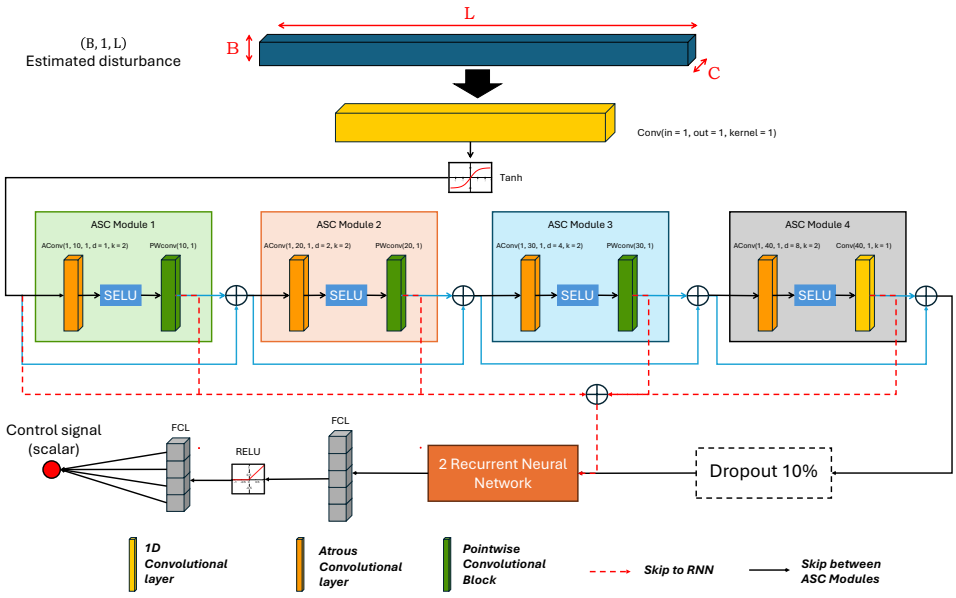


Figure 4.3: Architecture of the proposed MDNoiseNet. The network is composed of 1D causal, pointwise, and atrous convolutions to extract effective features from raw time-domain acoustic signals. The hyperparameters of each layer are specified in the form (input channels, output channels, kernel size, dilation). For layers without dilation, the dilation factor is denoted as d when applicable.

4.2.3. Training strategy

The proposed MDNoiseNet is trained using 10,000 one-second acoustic signals derived from the dataset, normalized to $[-1, 1]$. The dataset is split into 80% training and 20% validation, with mini-batches of 500 samples.

Optimization is performed using the Adam optimizer with a learning rate of 1×10^{-5}. A step learning rate scheduler (step size 10, decay factor 0.5) is applied. Weights are initialized using a zero-mean Gaussian distribution (std = 1), while biases are initialized to zero. Training is conducted for 100 epochs.

At each iteration, the model processes the last 100 samples of the estimated disturbance. The loss function combines three components: (i) mean squared error (MSE) between the predicted anti-noise and the target signal, (ii) an output energy penalty to limit excessive control effort, and (iii) an L2 regularization applied to bias parameters:

$$\mathcal{L} = \mathrm{MSE}(y_{\mathrm{anti}}, y_{\mathrm{target}}) + \lambda_{\mathrm{out}}\, y_{\mathrm{out}}^2 + \lambda_{\mathrm{L2}} \sum_{i \in \mathcal{B}} \|b_i\|_2^2, \tag{4.6}$$

where $\lambda_{\text{out}} = 1 \times 10^{-4}$ and $\lambda_{\text{L2}} = 1 \times 10^{-5}$ control the contribution of each regularization term. Here, y_{anti} denotes the predicted anti-noise, y_{target} the desired cancellation signal, $\mathcal{B}$ the set of bias parameters, and b_i the i-th bias term. The bias regularization encourages small bias values, improving generalization and reducing overfitting.

During training, a perfect secondary path model is assumed, corresponding to the impulse response of the empty vehicle configuration.

4.3. Deep transfer reinforcement learning active noise control

In this framework, the active noise control system is formulated as a reinforcement learning problem, where the vehicle cabin represents the environment and the FxMDNoiseNet-DTRL agent acts as the controller.

At each time step n, the agent observes a state s_n derived from the estimated disturbance and generates a control action $y(n) = \pi_\phi(s_n)$. This signal is applied through the secondary path to produce the anti-noise, and the resulting residual error $e(n)$ is used to compute the reward $r(n) = -|e(n)|$. By maximizing cumulative reward, the agent progressively learns an optimal control policy for noise attenuation.

FxMDNoiseNet-DTRL control framework is illustrated in Figure 4.4.

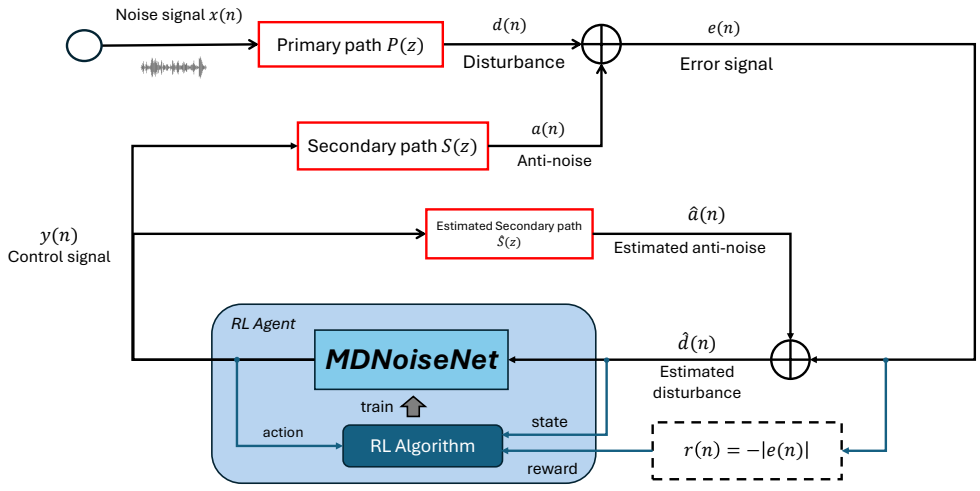


Figure 4.4: FxMDNoiseNet-DTRL control framework. The agent observes the state s_n (estimated disturbance) and generates the control action a_n. The residual error is used to compute the reward r_n, which guides policy learning.

The stability of the proposed deep-learning-based feedback controller was assessed primarily from an experimental perspective. The MDNoiseNet architecture uses causal processing and a bounded Tanh activation in its input stage, while the training procedure incorporates output-energy regularization and gradient clipping to limit excessive control actions and unstable parameter updates. For the reinforcement-learning-based controller, the use of target networks, delayed policy updates, gradient clipping, and ℓ_2 regularization further contributes to stable training. During deployment, the trained policy is kept fixed and no online parameter updates are performed. Across the experimental scenarios considered in this thesis, no divergence or unbounded growth of the residual error was observed. A formal Lyapunov- or input-to-state-stability analysis of the complete nonlinear feedback system is beyond the scope of this work and is left for future research.

4.3.1. Twin Delayed Deep Deterministic Policy Gradient

4.3.1.1. Implementation of the algorithm

Twin Delayed Deep Deterministic Policy Gradient (TD3) [65] is an off-policy actor–critic reinforcement learning algorithm for continuous control. It extends Deep Deterministic Policy Gradient (DDPG) [66] by addressing two main issues in deterministic actor–critic methods: overestimation bias in value functions and high variance in policy updates.

Overestimation arises from function approximation errors in action–value learning. In value-based RL, maximization over noisy estimates leads to systematically overestimated targets [67]. TD3 further shows that deterministic actor–critic methods also suffer from this issue, since the policy is optimized using an approximate critic, causing error propagation and reinforcement of suboptimal actions [65].

To mitigate this, TD3 uses Clipped Double Q-learning with two critics Q_{θ_1} and Q_{θ_2}. The target value is computed as:

$$y_n = r_n + \gamma(1 - d_n) \min_{i=1,2} Q_{\theta'_i}(s_{n+1}, \tilde{a}_{n+1}), \tag{4.7}$$

where γ is the discount factor and d_n indicates episode termination. Taking the minimum reduces overestimation, at the cost of slight underestimation, which is preferable since it is not amplified during policy updates [65, 68].

Each critic is trained using the Bellman error with ℓ_2 regularization:

$$\mathcal{L}_Q = \mathbb{E}\Big[(Q_{\theta_1}(s_n, a_n) - y_n)^2 + (Q_{\theta_2}(s_n, a_n) - y_n)^2\Big] + \lambda_Q \left(\|\theta_1\|_2^2 + \|\theta_2\|_2^2\right), \tag{4.8}$$

with λ_Q controlling regularization strength.

To improve robustness, TD3 applies target policy smoothing by adding clipped Gaussian noise:

$$\tilde{a}_{n+1} = \pi_{\phi'}(s_{n+1}) + \epsilon, \quad \epsilon \sim \text{clip}(\mathcal{N}(0, \sigma), -c, c). \tag{4.9}$$

This reduces sensitivity to sharp value peaks and is related to Expected SARSA [27].

Exploration is ensured by adding noise during interaction with the environment:

$$a_n = \pi_\phi(s_n) + N_{\text{ex},n}, \tag{4.10}$$

where $N_{\text{ex},n}$ is typically Gaussian and optionally clipped. In this work, this noise is adapted to the ANC task. It improves exploration of state–action space in continuous control settings [65, 66].

TD3 further stabilizes learning by decoupling policy and value updates: the critic is updated at every step, while the actor is updated every d iterations. This delayed update improves gradient reliability by allowing critics to converge more before influencing the policy [**?**].

Stability is further improved using Polyak-averaged target networks:

$$\theta_i' \leftarrow \tau\theta_i + (1-\tau)\theta_i', \quad \phi' \leftarrow \tau\phi + (1-\tau)\phi', \tag{4.11}$$

with $\tau \ll 1$, reducing temporal-difference noise accumulation [69].

The actor $\pi_\phi : \mathcal{S} \to \mathcal{A}$ maps states to actions $a_n = \pi_\phi(s_n)$ and is optimized by maximizing the first critic:

$$\mathcal{L}_\pi = -\mathbb{E}_{s_n \sim \mathcal{D}}\left[Q_{\theta_1}(s_n, \pi_\phi(s_n))\right] + \lambda_\pi \|\phi\|_2^2, \tag{4.12}$$

where $\mathcal{D}$ is the replay buffer and λ_π controls regularization.

Using a single critic for the actor update reduces complexity while preserving double Q-learning benefits.

The actor and critics are implemented as neural networks adapted to the control task, as illustrated in Figure 4.5.

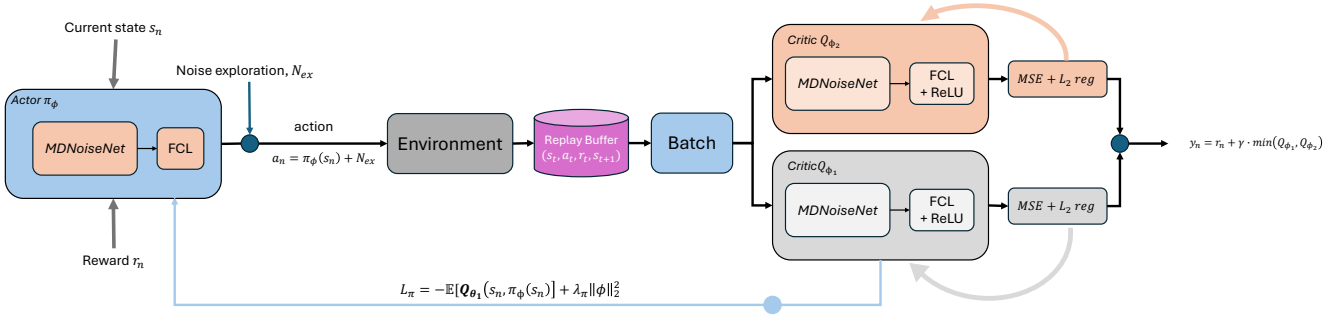


Figure 4.5: Actor and critic parameters updates.

4.3.1.2. Training Strategy

The training strategy is designed to adapt TD3 to real-world active noise control (ANC) with engine noise, improving stability, robustness, and sample efficiency beyond the standard framework.

To improve exploration, a frequency-adaptive strategy is introduced instead of standard Gaussian noise. At each step, the dominant frequency of the residual error signal is estimated using FFT, and the exploration frequency is sampled around it:

$$f_{\text{ex}} = f_{\text{dom}} + \mathcal{U}(-\Delta f, \Delta f), \tag{4.13}$$

where Δf defines the exploration bandwidth.

The exploration noise added to the deterministic action $a_n = \pi_\phi(s_n)$ is defined as:

$$\boldsymbol{N}_{\text{ex},n} = \rho_n \begin{bmatrix} \cos(2\pi f_{\text{ex}} n + \varphi_n) \\ \sin(2\pi f_{\text{ex}} n + \varphi_n) \end{bmatrix}, \tag{4.14}$$

with $\varphi_n \sim \mathcal{U}(-\pi, \pi)$ and:

$$\rho_n = \rho_{\text{ex}} \hat{A}_e, \tag{4.15}$$

ensuring amplitude-adaptive and physically meaningful exploration.

A warm-up phase with uniformly random actions is first used to populate the replay buffer before policy-driven exploration begins.

To stabilize training, the ASC modules are initially frozen and progressively activated. Learning starts at $t_{\mathrm{ASC}} = 90{,}000$ steps, with linear unfreezing over $\Delta t_{\mathrm{ASC}} = 200{,}000$ steps and a maximum learning rate of $\alpha_{\mathrm{ASC}}^{\max} = 1 \times 10^{-6}$. This avoids catastrophic forgetting while preserving pretrained representations.

ℓ_2 regularization and gradient clipping are applied to both actor and critic networks to ensure stable optimization under strong actor–critic coupling. In addition, target action perturbations use the same frequency-adaptive noise to improve robustness of value estimation.

The reward is defined directly from the residual error, $r_n = -|e_n|$ which directly encourages noise attenuation. Its evolution is analyzed during training.

Table 4.1: TD3 training hyperparameters used for the ANC task.

Hyperparameter	**Value**
Actor learning rate (α_π)	2×10^{-5}
Critic learning rate (α_Q)	1×10^{-4}
Discount factor (γ)	0.2
Soft update coefficient (τ)	0.005
Policy delay	3
Replay buffer size	1×10^6
Batch size	512
ℓ_2 regularization (λ_π, λ_Q)	1×10^{-4}
Warm-up steps	60,000
Total training steps	900,000
Unfreeze start for ASC modules (Critic)	90,000 steps
ASC modules max learning rate ($\alpha_{\mathrm{ASC}}^{\max}$)	1×10^{-6}

4.4. Results

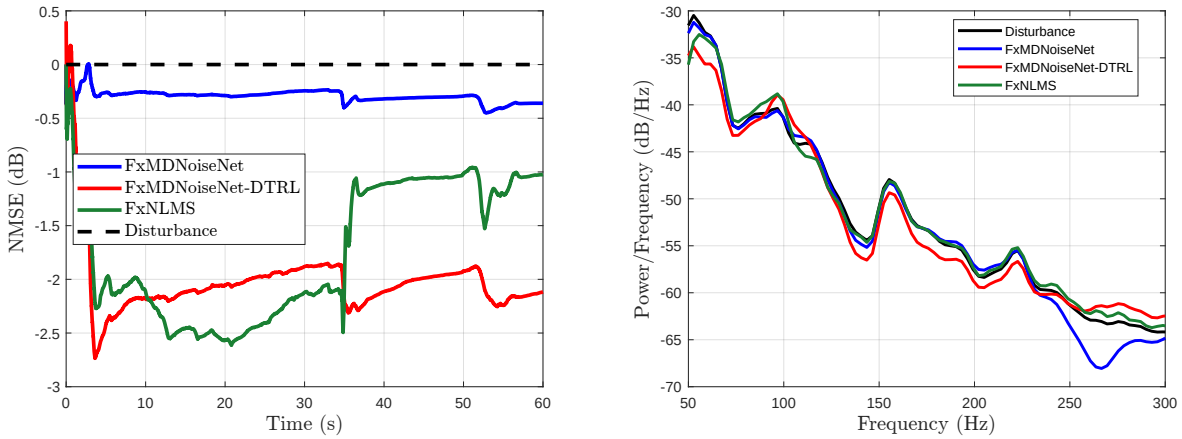
To evaluate robustness under non-stationary conditions, a linear time-varying (LTV) secondary path is introduced. At each time step, the true impulse response is randomly shifted using a circular shift:

$$\mathbf{s}_n = \mathcal{T}_{\Delta_n}(\mathbf{s}), \quad \Delta_n \in \{-1, 0, 1\}, \tag{4.16}$$

corresponding to a one-sample delay, advance, or no change. The estimated secondary path remains fixed, creating a model mismatch that emulates realistic operating conditions.

This mismatch combines two complementary sources of variability. First, the estimated and true secondary paths correspond to two genuinely different, experimentally measured vehicle configurations, the empty cabin and the cabin with two occupants, respectively, capturing the real acoustic loading effect of passengers. Second, the random ±1-sample circular shift is applied on top of this physical mismatch to emulate finer-grained, moment-to-moment non-stationary variation, for instance due to minor structural vibration, temperature drift, or road-induced changes, for which no dedicated physical measurement was available in this study. This combined test is therefore intended as a broad, worst-case robustness scenario, layering a measured, cause-specific mismatch with a generic stress-test perturbation, rather than being purely arbitrary. Deriving the shift magnitude and statistics directly from additional physical sources, such as temperature-driven changes in the speed of sound, is identified as a direction for future work.

Figure 4.6a shows that FxMDNoiseNet-DTRL achieves the best overall noise reduction in the time domain, while FxMDNoiseNet provides stable but lower attenuation. FxNLMS also achieves competitive reductions under these conditions.



(a) (b)

Figure 4.6: Comparison of (a) NMSE over time and (b) PSD for FxNLMS (green), FxMDNoiseNet (blue), and FxMDNoiseNet-DTRL (red) under LTV conditions with combined engine and road noise. The estimated secondary path corresponds to the empty vehicle, while the true path corresponds to the vehicle with two occupants.

In the frequency domain, FxMDNoiseNet-DTRL consistently achieves the strongest attenuation in the low-frequency band (50–250 Hz), where engine noise is dominant. This confirms its robustness in targeting the most energetic acoustic components, despite LTV-induced modeling errors.

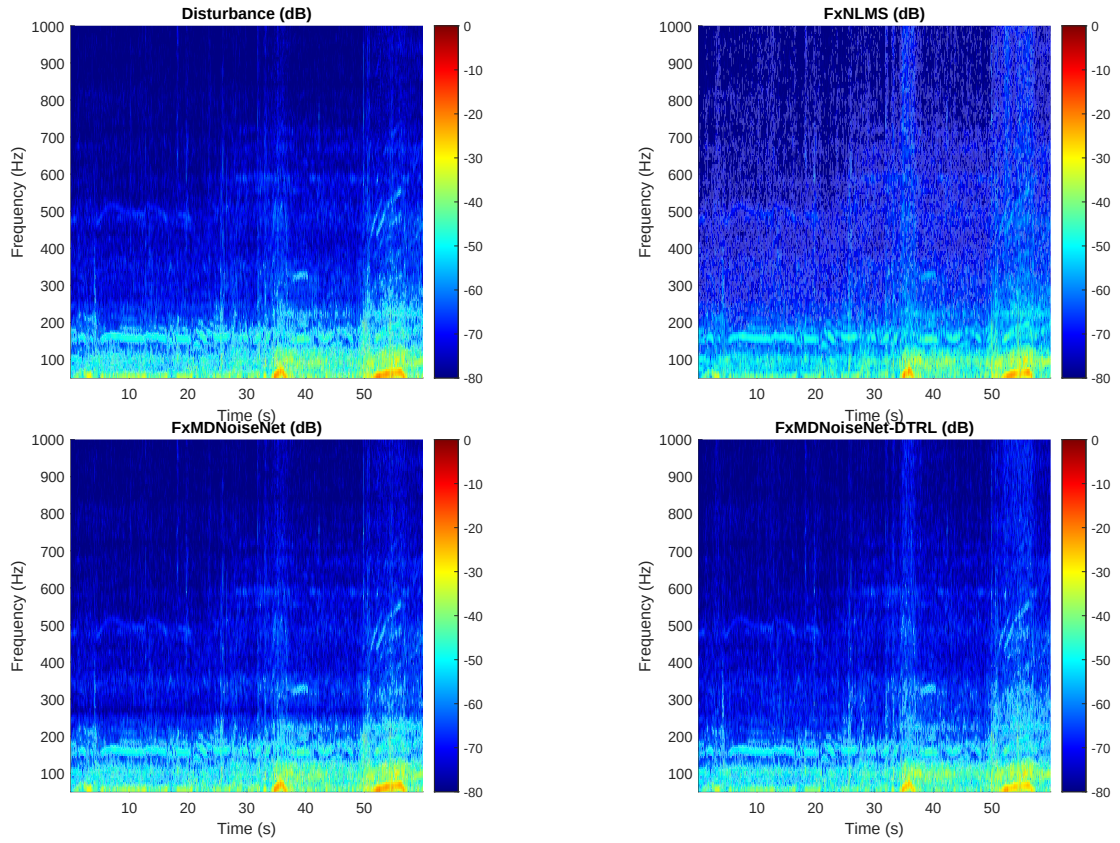


Figure 4.7: Spectrogram comparison of primary noise and error signals obtained with FxNLMS, FxMDNoiseNet, and FxMDNoiseNet-DTRL under LTV conditions.

Overall, FxMDNoiseNet-DTRL achieves the highest global noise reduction (2.2 dB), outperforming FxMDNoiseNet (0.3 dB) and FxNLMS (0.5 dB).

5 | Conclusions

5.1. Basic framework and motivations of appended papers

The appended papers of this thesis are built around a common objective: improving the robustness, adaptability, and perceptual effectiveness of active noise control (ANC) systems in complex vibro-acoustic environments. These studies progressively address key limitations of classical ANC methods, ranging from physical modeling challenges to nonlinearities, human perception, and learning-based control strategies.

A first line of work considers the physical limitations of ANC in complex structures. In particular, Paper A highlights the difficulties induced by non-minimum phase (NMP) dynamics in encapsulated vibro-acoustic systems. By relying on FEM-based modeling and state-space identification, it shows that classical adaptive approaches such as FxNLMS are fundamentally limited in such settings. To overcome these constraints, a Kalman filter-based formulation is introduced, significantly improving performance and demonstrating the importance of accurate system representation and robust estimation in challenging acoustic environments.

Beyond objective performance, Paper B extends the analysis to subjective perception of ANC. Using a controlled noise box simulating a vehicle interior, it demonstrates that perceived annoyance and loudness are strongly influenced by the type of noise and the ANC strategy. The results highlight that improved algorithmic performance does not always translate linearly into perceived comfort, emphasizing the importance of human-centered evaluation metrics in encapsulated environments.

Building on these foundations, Paper C addresses nonlinear distortions and computational constraints in ANC systems. It introduces a hybrid deep learning and adaptive filtering approach, where a CNN dynamically selects control sub-filters within a GFANC framework, combined with a THFxNLMS adaptation stage. This hybrid design improves robustness against nonlinearities, particularly those introduced by loudspeakers, while

maintaining computational efficiency and adaptability to varying acoustic conditions.

The final two papers shift the focus toward data-driven and learning-based control strategies for automotive ANC. Papers D and E investigate deep reinforcement learning (RL) combined with transfer learning to enhance adaptability in time-varying and nonlinear environments. Both works build upon the FxMDNoiseNet architecture within a filtered-x framework to explicitly handle secondary path effects. Paper D integrates a DDPG-based strategy with transfer learning to accelerate convergence, while Paper E further extends this idea using TD3 to improve stability and performance under more complex conditions, including secondary path mismatch and LTV dynamics. In both cases, real-world in-vehicle data and experimental impulse responses are used to ensure realistic validation.

5.2. Summary of paper A

This paper studies active broadband noise control in an encapsulated vibro-acoustic structure exhibiting non-minimum phase (NMP) behavior. The system is modeled using the finite element method (FEM) to capture multi-physical interactions and derive accurate state-space representations of the primary and secondary paths.

The analysis shows that the NMP nature of the system severely limits the performance of classical ANC methods such as FxNLMS, leading to reduced attenuation and transient distortions even under accurate modeling. This is mainly due to phase inversion and delay effects that cannot be fully compensated by standard adaptive approaches.

To overcome these limitations, a Kalman filter-based ANC strategy is proposed. By improving state estimation and robustness to NMP dynamics, it significantly enhances noise attenuation, achieving up to 16.92 dB reduction in the 0–450 Hz band compared to 2.72 dB for FxNLMS.

Overall, the paper contributes (i) a FEM-based NMP modeling of an encapsulated structure, (ii) state-space identification of control paths, and (iii) a Kalman-based ANC framework that clearly outperforms classical FxLMS methods.

5.3. Summary of paper B

This study investigates the subjective perception of active noise control (ANC) performance in a simulated vehicle interior using a noise box environment. Unlike traditional ANC evaluations based on objective metrics, the work focuses on human perception of noise reduction effectiveness.

Two algorithms are compared: the conventional NLMS and a hybrid selective fixed-filter ANC approach (SFANC-NLMS). Participants were exposed to different combinations of noise types (motorcycle, street, train), sound pressure levels (55, 65, and 72 dB(A)), and ANC conditions (no control, NLMS, SFANC-NLMS). Subjective evaluations were collected in terms of perceived annoyance (PAY), affective quality (PAQ), and perceived loudness (PLN).

The results show that SFANC-NLMS consistently outperforms NLMS in reducing perceived annoyance and loudness. While higher sound levels increase perceived annoyance, they do not significantly alter the relative performance of the ANC algorithms. Noise type also plays a role, with SFANC-NLMS providing more pronounced perceptual improvements across different acoustic scenarios.

Overall, the study highlights that SFANC-NLMS offers improved subjective noise reduction performance in encapsulated environments compared to classical NLMS-based ANC.

5.4. Summary of paper C

This paper proposes a hybrid active noise control (ANC) algorithm, GFANC-THFxNLMS, designed to improve robustness in the presence of nonlinear distortions, particularly those introduced by loudspeakers in encapsulated structures with openings. The method builds upon the GFANC framework, which decomposes a pre-trained broadband control filter into multiple sub-filters, enabling more flexible filter selection.

A lightweight 1D CNN is used to dynamically select and combine sub-control filters based on the characteristics of the incoming noise, while an adaptive THFxNLMS stage further refines the control filter in real time. This hybrid structure reduces computational complexity while maintaining adaptability in both time-invariant and time-varying conditions.

Experimental results show that GFANC-THFxNLMS outperforms conventional approaches, achieving 9.49 dB noise reduction compared to 8.01 dB for GFANC and 6.74 dB for THFxNLMS, demonstrating improved performance in complex acoustic environments.

Overall, the study highlights the effectiveness of combining deep learning-based filter selection with adaptive filtering to mitigate nonlinear effects in ANC systems.

5.5. Summary of paper D

This paper proposes an adaptive active noise control (ANC) framework for vehicle interiors that combines deep learning, reinforcement learning, and transfer learning. The approach

is based on FxMDNoiseNet, a modified deep neural network operating within a filtered-x structure to account for secondary path effects.

To improve adaptability in dynamic acoustic environments, a Deep Deterministic Policy Gradient (DDPG) reinforcement learning strategy is integrated into the control loop. The RL policy is initialized using pre-trained weights from FxMDNoiseNet, enabling transfer learning that accelerates convergence and improves stability.

The method is evaluated in offline simulations using real in-vehicle noise data, showing improved adaptability to time-varying acoustic conditions and enhanced noise attenuation compared to the standalone deep learning approach.

Overall, the study demonstrates the effectiveness of combining deep learning-based ANC with reinforcement and transfer learning for robust control in automotive environments.

5.6. Summary of paper E

This paper proposes a deep transfer reinforcement learning-based active noise control (ANC) framework for automotive environments with nonlinear and time-varying acoustic characteristics. The method builds upon FxMDNoiseNet, a neural network controller operating within a filtered-x architecture designed to handle secondary path effects more effectively.

To enhance adaptability, a Twin Delayed Deep Deterministic Policy Gradient (TD3) reinforcement learning strategy is integrated to fine-tune the control policy. Transfer learning is used by initializing the reinforcement learning agent with pre-trained weights from FxMDNoiseNet, which accelerates convergence and improves training efficiency.

The approach is evaluated in a realistic automotive scenario using experimentally measured impulse responses for both primary and secondary paths, ensuring realistic in-vehicle acoustic dynamics. Results show that the proposed FxMDNoiseNet-DTRL method outperforms both the standalone deep learning controller and classical FxNLMS, achieving improved noise attenuation under LTI and LTV conditions with secondary path mismatch.

Overall, the study demonstrates the benefit of combining filtered-x deep learning architectures with reinforcement learning and transfer learning for robust ANC in real vehicle environments.

5.7. Future work

First, the current implementation operates in a sample-by-sample manner, producing one control action per time step. While this allows fine temporal resolution, it may limit the exploitation of longer temporal dependencies in highly dynamic acoustic signals. A promising extension would be to develop a block-based formulation, where sequences of samples are processed jointly. This could improve temporal consistency, enhance stability, and potentially reduce computational overhead through more efficient parallel processing.

Second, the reinforcement learning component of the framework is based on the TD3 algorithm. Although TD3 provides improved stability compared to earlier actor–critic methods, other advanced continuous control algorithms could be investigated. In particular, distributional or multi-step variants such as D4PG, or more recent off-policy reinforcement learning methods, may further enhance convergence speed and policy robustness in highly non-stationary acoustic environments.

Finally, an important step toward practical deployment is real-time implementation of the proposed framework on embedded hardware. This would allow evaluation under strict computational constraints and real acoustic conditions, including latency effects, sensor noise, and hardware-induced nonlinearities. Such an implementation would provide a more complete validation of the approach and pave the way for integration into real vehicle ANC systems.

Bibliography

[1] R. H. Cabell and C. Fuller, "A principal component algorithm for feedforward active noise and vibration control," *Journal of Sound and Vibration*, vol. 227, no. 1, pp. 159–181, 1999.

[2] G. C. Goodwin, E. I. Silva, and D. E. Quevedo, "Analysis and design of networked control systems using the additive noise model methodology," *Asian Journal of Control*, vol. 12, no. 4, pp. 443–459, 2010.

[3] W. M. Hartmann, *Signals, sound, and sensation*. Springer Science & Business Media, 2004.

[4] Y. Li, X. Wang, R. Huang, and Z. Qiu, "Active vibration and noise control of vibro-acoustic system by using pid controller," *Journal of Sound and Vibration*, vol. 348, pp. 57–70, 2015.

[5] C. Fuller, C. Hansen, and S. Snyder, "Active control of sound radiation from a vibrating rectangular panel by sound sources and vibration inputs: an experimental comparison," *Journal of Sound and Vibration*, vol. 145, no. 2, pp. 195–215, 1991.

[6] M. Misol, S. Algermissen, and H. P. Monner, "Experimental investigation of different active noise control concepts applied to a passenger car equipped with an active windshield," *Journal of Sound and Vibration*, vol. 331, no. 10, pp. 2209–2219, 2012.

[7] P. Shorter, "Recent advances in automotive interior noise prediction," *SAE BRASIL Noise and Vibration Conference*, 2008.

[8] W. Belgacem, A. Berry, and P. Masson, "Active vibration control on a quarter-car for cancellation of road noise disturbance," *Journal of Sound and Vibration*, vol. 331, no. 14, pp. 3240–3254, 2012.

[9] P. A. NELSON and S. J. ELLIOTT, "Active noise control: A tutorial review," *IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences*, vol. 75, no. 11, pp. 1541–1554, 1992.

[10] F. Fahy and J. Walker, *Advanced applications in acoustics, noise and vibration*. CRC Press, 2018.

[11] S. J. Elliot, "Active control of nonlinear systems," *Noise Control Engineering Journal*, vol. 49, no. 1, pp. 30–53, 2001.

[12] J. Cheer and S. J. Elliott, "Multichannel control systems for the attenuation of interior road noise in vehicles," *Mechanical Systems and Signal Processing*, vol. 60, pp. 753–769, 2015.

[13] C. W. Isaac, S. Wrona, M. Pawelczyk, *et al.*, "Modelling vibro–acoustic response of lightweight square aluminium panel influenced by sound source locations for active control," *Scientific Reports*, vol. 12, no. 1, p. 10727, 2022.

[14] S. Wrona, M. Pawelczyk, and L. Cheng, "Sound transmission through a panel with a hybrid active and semi-active control system," *Journal of Sound and Vibration*, vol. 536, p. 117172, 2022.

[15] P. N. Samarasinghe, W. Zhang, and T. D. Abhayapala, "Recent advances in active noise control inside automobile cabins: Toward quieter cars," *IEEE Signal Processing Magazine*, vol. 33, no. 6, pp. 61–73, 2016.

[16] D. Shi, W.-S. Gan, B. Lam, and S. Wen, "Feedforward selective fixed-filter active noise control: Algorithm and implementation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 1479–1492, 2020.

[17] Z. Luo, D. Shi, and W.-S. Gan, "A hybrid sfanc-fxnlms algorithm for active noise control based on deep learning," *IEEE Signal Processing Letters*, vol. 29, pp. 1102–1106, 2022.

[18] Z. Luo, D. Shi, X. Shen, J. Ji, and W.-S. Gan, "Deep generative fixed-filter active noise control," in *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1–5, IEEE, 2023.

[19] M. Vetterli and D. Le Gall, "Perfect reconstruction fir filter banks: Some properties and factorizations," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 7, pp. 1057–1071, 1989.

[20] H. Zhang and D. Wang, "Deep anc: A deep learning approach to active noise control," *Neural Networks*, vol. 141, pp. 1–10, 2021.

[21] A. Mostafavi and Y.-J. Cha, "Deep learning-based active noise control on construction sites," *Automation in Construction*, vol. 151, p. 104885, 2023.

[22] Y.-J. Cha, A. Mostafavi, and S. S. Benipal, "Dnoisenet: Deep learning-based feedback active noise control in various noisy environments," *Engineering Applications of Artificial Intelligence*, vol. 121, p. 105971, 2023.

[23] Z. Luo, H. Ma, D. Shi, and W.-S. Gan, "Gfanc-rl: Reinforcement learning-based generative fixed-filter active noise control," *Neural Networks*, vol. 180, p. 106687, 2024.

[24] S. Ryu, J. Lim, and Y.-S. Lee, "Narrowband active noise control with ddpg based on reinforcement learning," *International Journal of Automotive Technology*, vol. 25, no. 6, pp. 1389–1397, 2024.

[25] Z. Cao, K. Jiang, W. Zhou, S. Xu, H. Peng, and D. Yang, "Continuous improvement of self-driving cars using dynamic confidence-aware reinforcement learning," *Nature Machine Intelligence*, vol. 5, no. 2, pp. 145–158, 2023.

[26] B. Raeisy and S. G. Haghighi, "Active noise controller with reinforcement learning," 2012. Proceedings of the 16th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP 2012).

[27] R. S. Sutton, A. G. Barto, *et al.*, *Reinforcement learning: An introduction*, vol. 1. MIT press Cambridge, 1998.

[28] COMSOL AB, "COMSOL Multiphysics®." `https://www.comsol.com`.

[29] L. Liu, F. Ripamonti, R. Corradi, and Z. Rao, "On the experimental vibroacoustic modal analysis of a plate-cavity system," *Mechanical Systems and Signal Processing*, vol. 180, p. 109459, 2022.

[30] A. Aboutiman, R. Shams, H. R. Karimi, F. Ripamonti, and M. Pawełczyk, "Active noise control in encapsulated structures with non-minimum phase characteristics using a kalman filter approach," *Journal of Sound and Vibration*, vol. 615, p. 119187, 2025.

[31] L. Cremer and M. Heckl, *Structure-borne sound: structural vibrations and sound radiation at audio frequencies*. Springer Science & Business Media, 2013.

[32] D. Chapelle and K.-J. Bathe, *The finite element analysis of shells-fundamentals*. Springer Science & Business Media, 2010.

[33] K. S. A. Maamoun, S. Wrona, M. Pawelczyk, and H. R. Karimi, "Shaping of the frequency response of vibrating plates with openings for vibro-acoustic systems," *Mechanical Systems and Signal Processing*, vol. 218, p. 111539, 2024.

[34] P. Lenk and G. Coult, "Damping of glass structures and components," in *Challenging Glass Conference Proceedings*, vol. 2, pp. 341–350, 2010.

[35] J. T. Du, W. L. Li, Z. G. Liu, H. A. Xu, and Z. L. Ji, "Acoustic analysis of a rectangular cavity with general impedance boundary conditions," *The Journal of the Acoustical Society of America*, vol. 130, no. 2, pp. 807–817, 2011.

[36] P. Morse, "Sound wave in rooms," *Rev. Mod. Phys.*, vol. 16, pp. 100–117, 1944.

[37] R. J. Astley, G. J. Macaulay, and J.-P. Coyette, "Mapped wave envelope elements for acoustical radiation and scattering," *Journal of Sound and Vibration*, vol. 170, no. 1, pp. 97–118, 1994.

[38] A. Valera-Rivera and A. E. Engin, "Aaa algorithm for rational transfer function approximation with stable poles," *IEEE Letters on Electromagnetic Compatibility Practice and Applications*, vol. 3, no. 3, pp. 92–95, 2021.

[39] I. V. Gosea and S. Gugercin, "Data-driven modeling of linear dynamical systems with quadratic output in the aaa framework," *Journal of Scientific Computing*, vol. 91, no. 1, p. 16, 2022.

[40] P. A. Lopes and M. S. Piedade, "A Kalman filter approach to active noise control," in *2000 10th European Signal Processing Conference (EUSIPCO)*, pp. 1–4, IEEE, 2000.

[41] S. van Ophem and A. P. Berkhoff, "Multi-channel Kalman filters for active noise control," *Journal of the Acoustical Society of America*, vol. 133, no. 4, pp. 2105–2115, 2013.

[42] J. Fabry, S. Liebich, P. Vary, and P. Jax, "Active noise control with reduced-complexity Kalman filter," in *2018 16th International Workshop on Acoustic Signal Enhancement (IWAENC)*, pp. 166–170, IEEE, 2018.

[43] D. Shi, B. Lam, K. Ooi, X. Shen, and W.-S. Gan, "Selective fixed-filter active noise control based on convolutional neural network," *Signal Processing*, vol. 190, p. 108317, 2022.

[44] W. Dai, C. Dai, S. Qu, J. Li, and S. Das, "Very deep convolutional neural networks for raw waveforms," in *2017 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pp. 421–425, IEEE, 2017.

[45] M. Cartwright, J. Cramer, A. E. M. Mendez, Y. Wang, H.-H. Wu, V. Lostanlen, M. Fuentes, G. Dove, C. Mydlarz, J. Salamon, *et al.*, "Sonyc-ust-v2: An urban sound tagging dataset with spatiotemporal context," *arXiv preprint arXiv:2009.05188*, 2020.

[46] A. Aboutiman, Z. Rachman, T. Oberman, F. Aletta, J. Kang, H. R. Karimi, and F. Ripamonti, "Subjective perception analysis of active noise control algorithms in an encapsulated structure: An experimental study," *Applied Acoustics*, vol. 239, p. 110823, 2025.

[47] F. Aletta, T. Vander Mynsbrugge, D. Van de Velde, P. De Vriendt, P. Thomas, K. Filipan, D. Botteldooren, and P. Devos, "Awareness of 'sound'in nursing homes: A large-scale soundscape survey in flanders (belgium)," *Building Acoustics*, vol. 25, no. 1, pp. 43–59, 2018.

[48] I. O. for Standardization, "Iso/ts 15666 acoustics — assessment of noise annoyance by means of social and socio-acoustic surveys (2nd ed.)," Technical Specification ISO/TS 15666, International Organization for Standardization, Geneva, Switzerland, 2021.

[49] F. Karam, A. Bérard, O. Sheehy, M.-C. Huneau, G. Briggs, C. Chambers, A. Einarson, D. Johnson, K. Kao, G. Koren, *et al.*, "Reliability and validity of the 4-item perceived stress scale among pregnant women: Results from the otis antidepressants study," *Research in nursing & health*, vol. 35, no. 4, pp. 363–375, 2012.

[50] A. Mitchell, T. Oberman, F. Aletta, M. Erfanian, M. Kachlicka, M. Lionello, and J. Kang, "The soundscape indices (ssid) protocol: a method for urban soundscape surveys—questionnaires with acoustical and contextual information," *Applied Sciences*, vol. 10, no. 7, p. 2397, 2020.

[51] B. Lam, K. C. Q. Lim, K. Ooi, Z.-T. Ong, D. Shi, and W.-S. Gan, "Anti-noise window: Subjective perception of active noise reduction and effect of informational masking," *Sustainable Cities and Society*, vol. 97, p. 104763, 2023.

[52] I. O. for Standardization, "Acoustics — soundscape — part 2: Data collection and reporting requirements," tech. rep., ISO/TS, 2018. ISO/TS 12913-2:2018.

[53] ISO, "Iso 532-1. acoustics—methods for calculating loudness—part 1: Zwicker method," 2017.

[54] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika*, vol. 52, no. 3-4, pp. 591–611, 1965.

[55] H. Levene, "Robust tests for equality of variances," 1960.

[56] J. W. Mauchly, "Significance test for sphericity of a normal n-variate distribution," *The Annals of Mathematical Statistics*, vol. 11, no. 2, pp. 204–209, 1940.

[57] K. V. Mardia, "Measures of multivariate skewness and kurtosis with applications," *Biometrika*, vol. 57, no. 3, pp. 519–530, 1970.

[58] Analog Devices, "ADSP-SC598 Product Page," 2024. Accessed: 2024-12-29.

[59] K. S. A. Maamoun, S. Wrona, M. Pawelczyk, and H. R. Karimi, "Shaping of the frequency response of vibrating plates with openings for vibro-acoustic systems," *Mechanical Systems and Signal Processing*, vol. 218, p. 111539, 2024.

[60] A. Aboutiman, K. S. A. Maamoun, H. R. Karimi, and F. Ripamonti, "Hybrid deep learning–based active noise control for encapsulated structures with openings," *Expert Systems with Applications*, p. 131247, 2026.

[61] S. Ghasemi, R. Kamil, and M. H. Marhaban, "Nonlinear thf-fxlms algorithm for active noise control with loudspeaker nonlinearity," *Asian Journal of Control*, vol. 18, no. 2, pp. 502–513, 2016.

[62] H. H. Dam, S. Nordholm, A. Cantoni, and J. M. de Haan, "Iterative method for the design of dft filter bank," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 51, no. 11, pp. 581–586, 2004.

[63] O. J. Tobias and R. Seara, "On the lms algorithm with constant and variable leakage factor in a nonlinear environment," *IEEE transactions on signal processing*, vol. 54, no. 9, pp. 3448–3458, 2006.

[64] P. Balazs, B. Laback, and P. Majdak, "Multiple exponential sweep method for fast measurement of head related transfer functions," in *Audio Engineering Society Convention 122*, Audio Engineering Society, 2007.

[65] S. Fujimoto, H. van Hoof, and D. Meger, "Addressing function approximation error in actor-critic methods," *CoRR*, vol. abs/1802.09477, 2018.

[66] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.

[67] S. Thrun and A. Schwartz, "Issues in using function approximation for reinforcement learning," 1993. In: Proceedings of the Connectionist Models Summer School, pp. 255–263.

[68] H. Hasselt, "Double q-learning," *Advances in neural information processing systems*, vol. 23, 2010.

[69] V. Mnih, A. P. Badia, M. Mirza, A. Graves, T. Lillicrap, T. Harley, D. Silver, and K. Kavukcuoglu, "Asynchronous methods for deep reinforcement learning," 2016. In: International Conference on Machine Learning (ICML), pp. 1928–1937.

Attached Papers

Five papers listed below are attached in Appendix:

Paper A Alkahf Aboutiman, Reza Shams, Hamid Reza Karimi, Francesco Ripamonti and Marek Pawełczyk. "Active noise control in encapsulated structures with non-minimum phase characteristics using a Kalman filter approach" *Journal of Sound and Vibration*, 615 (2025) 119187.
https://doi.org/10.1016/j.jsv.2025.119187

Paper B Alkahf Aboutiman, Zulfi Rachman, Tin Oberman, Francesco Alletta, Jian Kang, Hamid Reza Karimi and Francesco Ripamonti. "Subjective perception analysis of active noise control algorithms in an encapsulated structure: An experimental study" *Applied Acoustics*, 239 (2025) 110823.
https://doi.org/10.1016/j.apacoust.2025.110823

Paper C Alkahf Aboutiman, Khaled Said Ahmed Maamoun, Hamid Reza Karimi and Francesco Ripamonti. "Hybrid deep learning–based active noise control for encapsulated structures with openings" *Expert Systems With Applications*, 309 (2026) 131247.
https://doi.org/10.1016/j.eswa.2026.131247

Paper D Alkahf Aboutiman, Jan Kirchhof and Francesco Ripamonti. "Transfer Learning-Enhanced DDPG for Adaptive Active Noise Control in Vehicle Interiors" *Thirty-Ninth Conference on Neural Information Processing Systems (NeurIPS 2026)*, 2026.
Under Review

Paper E Alkahf Aboutiman, Jan Kirchhof, Hamid Reza Karimi and Francesco Ripamonti. "Deep transfer reinforcement learning for active noise control in vehicle interiors" *Nature Machine Intelligence*, 2026.
Submitted

Journal of Sound and Vibration 615 (2025) 119187

Contents lists available at ScienceDirect

Journal of Sound and Vibration

journal homepage: www.elsevier.com/locate/jsvi

Active noise control in encapsulated structures with non-minimum phase characteristics using a Kalman filter approach

Alkahf Aboutiman [a,*], Reza Shams [a], Hamid Reza Karimi [a], Francesco Ripamonti [a], Marek Pawełczyk [b]

[a] Politecnico di Milano, Department of Mechanical Engineering, Milan, Italy
[b] Silesian University of Technology, Department of Measurements and Control Systems, Gliwice, Poland

ARTICLE INFO

Keywords:
Active noise control
Encapsulated structure
Non-minimum phase system
Finite element model

ABSTRACT

This paper studies the problem of active broadband noise control for an encapsulated structure. The structure is analyzed using the finite element method (FEM) to capture the interactions within the multi-physical domains of vibro-acoustic systems, allowing a detailed investigation of the propagation of sound waves and their interaction with the structural components. The frequency response function (FRF) of each control path, such as primary path and secondary path, is identified to quantify the dynamic behavior of the system across different frequencies. A complex identification method is used to calculate the state-space representations of each path for implementing an effective active noise control (ANC) algorithm. It is shown that the system under investigation exhibits non-minimum phase characteristics, a challenging aspect in ANC due to phase delays and inverted dynamics that complicate achieving precise noise cancellation. Traditional ANC algorithms, such as the Filtered-x Normalized Least Mean Squares (FxNLMS), struggle in dealing with non-minimum phase (NMP) characteristics, resulting in limited noise reduction with significant delays and undershoot effects. To overcome these limitations, this paper proposes the application of a Kalman filter approach in the ANC system, which offers enhanced efficiency and robustness in controlling the system under consideration. In comparison, the performance of the Kalman filter approach for noise reduction in the frequency band from 0 to 450 Hz can be quantified as 16.92 dB, compared to 2.72 dB for the FxNLMS approach.

1. Introduction

Active noise control (ANC) is a traditional method of mitigating acoustic, noise or vibration signals based on the principle of superposition. More specifically, ANC systems aim to create an "anti-noise" signal that matches the amplitude and has the opposite phase of the unwanted noise, effectively eliminating it [1,2]. Fundamentally, ANC requires predicting both the amplitude and phase of a noise signal at a given point in space ahead of time. While the signal amplitude may remain constant over time, the signal phase continuously changes at any spatial location due to the nature of acoustic waves [3,4]. ANC technologies can be found in headphones, automobiles, aviation, buildings, appliances, industrial machinery, medical equipment, and maritime applications, enhancing acoustic comfort and reducing noise pollution across various environments. Usually, in encapsulated structures such as cars, trains, or aircraft, passive methods are mainly used through traditional sound-absorbing materials. However, at low frequencies, passive techniques can be impractical due to their mass requirements [5–9]. This is where ANC can provide an effective solution,

* Corresponding author.
E-mail address: alkahfahmat.aboutiman@polimi.it (A. Aboutiman).

https://doi.org/10.1016/j.jsv.2025.119187
Received 13 October 2024; Received in revised form 24 April 2025; Accepted 4 May 2025
Available online 17 May 2025
0022-460X/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

reducing noise at these frequencies in such structures. The main sources of low-frequency noise in the cabin are primarily structural vibrations emanating from the engine and tires [10,11]. These include structure-borne noise, which is characterized by mechanical vibrations of the tires, most of which are below 1 kHz, and engine noise, the main frequencies of which are below 500 Hz. In addition, low-frequency noise tends to be amplified by resonances within the cabin structure. Road noise, caused by the interaction between the wheels and the road surface, is another important source of noise in automobiles [12].

Active control of interior road noise in cars has been studied for over two decades, leading to the development of various ANC techniques aimed at enhancing ride comfort [13,14]. In [15], an $H\infty$ algorithm is presented, achieving a reduction of 10 dB at 40 Hz for road noise reduction. [16] utilizes a headrest system to reduce interior road noise around the listener's ears using a remote microphone technique. In [17], an Active Structural Acoustic Control (ASAC) system with an FxLMS algorithm is applied in a car, resulting in a noise reduction of 4 dB between 20 Hz and 250 Hz. [18,19] demonstrate that a multi-input multi-output system is required to achieve a significant reduction in broadband noise. [20] highlights the use of a leaky-FxLMS algorithm in a van, controlling multi-harmonic noise with reductions from 15 dB to 20 dB up to 400 Hz. This work can also be done on numerical model. For example in [21] a Modified FxLMS (MFxLMS) algorithm is applied on a numerical model of vehicle interior. With this algorithm a reduction of 8 dB from 30 Hz to 145 Hz was obtained. Numerical tools have become popular in acoustical research and are frequently used to ANC study methods. However, evaluating the effectiveness of ANC in a complex three-dimensional space is challenging. Accurately modeling acoustics is a daunting task [22], primarily due to the complexity of noise sources and boundary impedances. The Finite Element Method (FEM) can simulate acoustical fields in ANC systems [23–26]. This method has also been combined with optimization techniques to determine optimal sensor and actuator locations, as well as the strengths of acoustic sources.

As we know, the performance of ANC systems heavily depends on the characteristics of the overall system [27]. To improve control results, high modeling accuracy of the system is required [28,29]. Kalman filtering has been explored in ANC for its improved adaptation and control performance. [30] introduced it as an alternative to LMS-based algorithms, [31] extended it to multi-channel ANC, and [32] proposed a reduced-complexity version for real-time applications. However, problems can arise when the system on which we want to apply control is a non-minimum phase (NMP) system. Specifically, if the secondary path is a non-minimum phase, it will create a waterbed effect on the control system, significantly impacting the controller design and deteriorating control performance [33,34], whether the controller is designed using FxLMS or $H\infty$ robust control algorithms [35]. In such a scenario, even perfect modeling cannot achieve satisfactory results [36]. Although the aforementioned adaptive control algorithms can reduce noise with time-varying primary paths, the best control results cannot be achieved due to their large gradient estimation deviation of the cost function for time-varying noise. In the ANC domain, there are some attempts to find a practical solution to overcome the NMP characteristics of the model. In [37] the problem is solved through the usage of a causal optimal controller to cancel the car road noise. The controller result with a reduction up to 12 dB between 30 Hz and 130 Hz. In [38] a repetitive control (RC) algorithm is developed to attenuate periodic impulsive noise. Results achieved a noise reduction of approximately 30 dB. The frequency range for the primary noise signal, which was filtered using a second-order low-pass filter, has a cutoff frequency of 400 Hz. In [39] a robust control method is proposed for non-minimum phase systems using a disturbance observer (DOB). The external controller is designed to meet performance requirements. The results show acceptable amplitude and phase responses, with the DOB's bandwidth set at 11 rad/s, analyzed over a frequency range from 10^{-2} to 10^{5} rad/s. The main contributions of this paper are threefold: (1) Developing a FEM model of an encapsulated structure with non-minimum phase characteristics, (2) Proposing an identification technique to estimate the state-space representation of the dynamic system for the reference path, primary path, and secondary path of the control algorithm, and (3) Utilizing a Kalman filter to overcome the limitations imposed by the FxNLMS algorithm on a non-minimum phase system.

The paper consists of 5 sections. Section 2 presents the characteristics of the FEM model used to simulate the environment on which the control is applied. In Section 3, the process used to compute the frequency response function for all control paths and the identification method are described. Section 4 outlines the two ANC algorithms applied in this environment, FxNLMS and the Kalman filter. Results and further discussion regarding the noise reduction achieved with both algorithms are provided in Section 5.

2. FEM modeling of a vibro-acoustic system

2.1. Model definition and geometry

For the finite element model, a cavity was chosen that is modeled on the interior of a car, the so-called Noise box. In order to simulate the sound field of the closed structure, it is preferable to avoid modeling the inner and outer fields for reasons of computational cost. Only the air cavity and the vibrating part of the plate are modeled here. The cavity has the shape of a pentagonal prism with an additional cavity at the position of the plate. The dimensions of the cavity are $0.825\,\mathrm{m} \times 1.018\,\mathrm{m} \times 0.645\,\mathrm{m}$.

A rectangular aluminum plate with dimensions $a \times b = 0.825\,\mathrm{m} \times 1.018\,\mathrm{m}$ was used at the opening of the cavity. The plate properties are presented in Table 1

Fig. 1 shows the dimensions of the noise box used in the model. COMSOL Multiphysics® was used for finite element modeling [40].

Table 1
Physical properties of the plate material.

Property	Value
Thickness, t	0.0004 m
Density, ρ	2700 kg m^{-3}
Modulus of elasticity, E	70×10^9 Pa
Poisson's ratio, ν	0.33

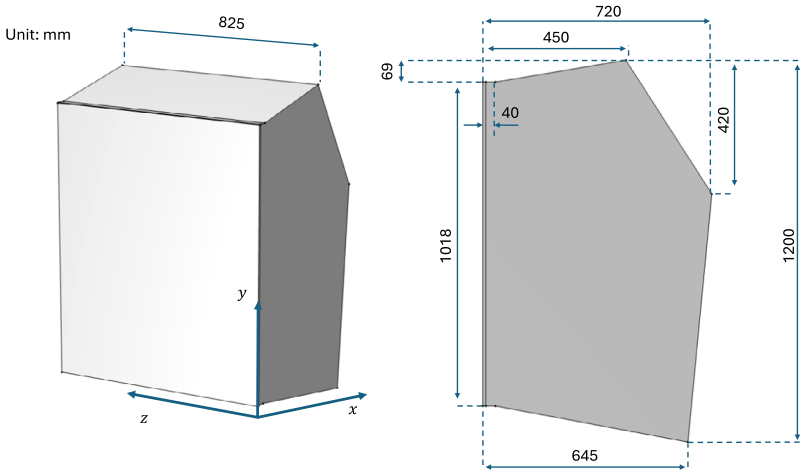


Fig. 1. Schematic of the Noise Box with dimensions used for the finite element model.

2.2. Physics

The aluminum plate was modeled in 3D using the Shell physics module. The shell elements are quadratic by default and follow the Mindlin-Reissner theory. The MITC (Mixed Interpolation of Tensorial Components) interpolation method was chosen, allowing the elements to model both thin shells (Kirchhoff theory) and thick shells (Mindlin theory). In this case, the plate is very thin, with a thickness-to-span ratio $h/a = 0.0048$, which is less than 1/200 and much smaller than 1/20. According to the literature, Kirchhoff's theory is applicable when this ratio is below 1/20 [41–43]. Therefore, the Kirchhoff–Love theory is most suitable for this study. The shell elements have three degrees of freedom (DOF) in the x, y, and z directions. Damping is defined with an isotropic structural loss factor $\eta = 0.003$, chosen based on median values for aluminum plates from [44]. The plate's edges are fixed to the cavity.

The cavity is modeled acoustically using the "Acoustics Frequency Domain" study. Assuming static pressure remains at 1 atm, air properties are defined by $\rho_0 = 1.2043$ kg/m^3 and $c = 343.2$ m/s at 20 °C. All surrounding walls are assumed to be acoustically rigid ($\overline{v}_n = 0$). Impedance is defined to simulate a reverberation room using $Z_0 = 100 \times \rho_0 \times c$. This impedance will be used to calculate the radiated impedance, Z_R, of the vibrating surfaces. The radiated impedance is the ratio of pressure on the vibrating surface (wall pressure) to the normal vibrational velocity of the surface, expressed as:

$$Z_R = \frac{\hat{p}(\omega)}{i\omega\hat{x}(\omega)} \tag{1}$$

where $\hat{p}$ is the wall pressure and $\hat{x}$ is the displacement of a cavity surface. This impedance is complex, with its real part being the resistance R_R and its imaginary part being the reactance X_R, as follows:

$$Z_R = R_R + iX_R \tag{2}$$

The radiated resistance R_R acts as a damping term, and the reactance X_R serves as an added mass term, both depending on Z_0 and impacting the frequency response amplitude and phase [45,46].

For meshing, triangular elements are used for the panel with a maximum size of 0.024 m, while tetrahedral elements are used for the air cavity, with a maximum size $\lambda_{\max}$ given by:

$$\lambda_{max} = \frac{c}{N \cdot f_{max}} \tag{3}$$

In the context of sound wave propagation in air, where $c = 343$ m/s and $f_{\max} = 1000$ Hz is the maximum frequency to be solved, N represents the number of mesh elements per wavelength. It is recommended to use a mesh with 7 to 10 elements per wavelength [47]. This vibroacoustic problem falls under multiphysics studies, requiring acoustic-structure coupling due to the interaction between the structural panel and the acoustic cavity.

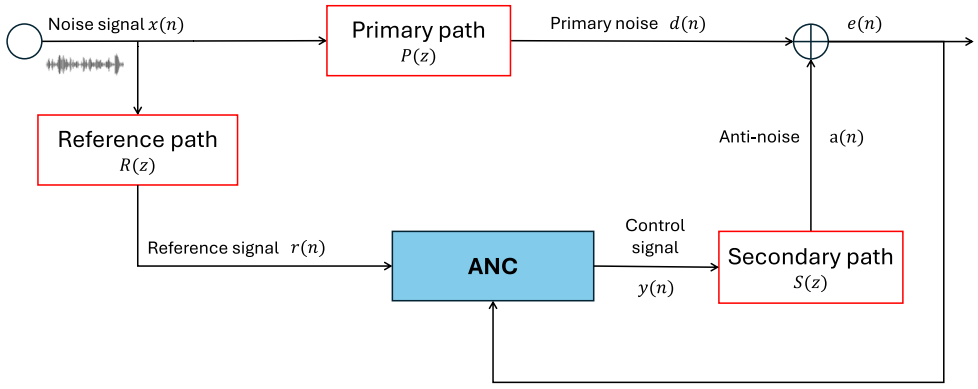


Fig. 2. Block diagram of the ANC algorithm. The control paths from which the FRFs were extracted are highlighted in red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3. Computation and identification of the vibro-acoustic system

The finite element model developed in previous section is used to compute the frequency response functions (FRFs) of all control paths in the active noise control (ANC) algorithm. These paths include the reference path, the primary path, and the secondary path in our case. Fig. 2 illustrates how these control paths are utilized in the control algorithm. These FRFs will enable us to establish the environment in which our control algorithm will operate.

3.1. Frequency response function computation

To obtain these paths, it is necessary to extract the frequency response of the Noise Box for the three paths. To achieve this, we defined the disturbance as a plane wave. In the model, the acoustic pressure corresponding to this wave type is applied to the aluminum plate. The calculation of this acoustic pressure considers the incident wave pressure on the plate, p_{inc}, and the pressure reflected by it, p_{refl}. The total acoustic pressure applied to the aluminum plate is the sum of these two pressures, denoted as p_{total}. These pressures are defined in Eqs. (4), (5), and (6).

$$p_{\text{inc}} = Ae^{i\phi}e^{-i(k_x x + k_y y + k_z z)}e^{i\Phi} \tag{4}$$

$$p_{\text{refl}} = Ae^{i\phi}e^{-i(-k_x x + k_y y + k_z z)}e^{i\Phi} \tag{5}$$

$$p_{\text{total}} = p_{\text{inc}} + p_{\text{refl}} \tag{6}$$

In this context, $k_x = k \cdot \sin(\theta) \cdot \sin(\phi)$, $k_y = k \cdot \sin(\theta) \cdot \cos(\phi)$, and $k_z = k \cdot \cos(\theta)$ are the wave numbers, while Φ represents the phase of the plane wave. Additionally, θ and ϕ are the angles describing the direction of the plane wave, as illustrated in Fig. 3.

Initially, the computation of the frequency responses is performed for a Single-Input-Single-Output (SISO) system. Note that during the simulation, each sensor is configured as a 'Point Probe' with the difference between them being the physical entity they will measure. In our case, there is one reference accelerometer, one error microphone and one canceling loudspeaker. The frequency response of the reference path is calculated as the ratio between the incident acoustic pressure induced by the plane wave, in Pascals, and the acceleration of the plate vibration in the normal direction to the plate plane, in m/s^2. In the end, the magnitude of the frequency response of the reference path is in $(\text{m/s}^2)/\text{Pa}$. For the primary path, the frequency response is calculated by taking the ratio between the incident acoustic pressure induced by the plane wave in Pascals and the pressure measured by the error microphone inside the Noise Box, also in Pascals. The amplitude of the frequency response of the primary path will be dimensionless. Regarding the secondary path, a monopole is used to model the canceling loudspeaker, with a volume rate flow equal to $1\,\text{m}^3/\text{s}$. Here, the amplitude S is defined in Eq. (7).

$$S = \frac{e^{i\phi} i\omega\rho}{4\pi} Q \tag{7}$$

Here, ρ represents the density, ϕ the phase, ω the angular frequency, and Q the volume rate flow of the source. Therefore, the frequency response for this path is the ratio between the monopole volume rate flow in m^3/s and the pressure in Pascals. The magnitude of the frequency response of the secondary path will be in $\text{Pa}/(\text{m}^3/\text{s})$. The frequency responses were computed from 10 Hz to 1000 Hz with a step of 1 Hz. The entire process is illustrated in Fig. 3 and the frequency response functions are presented in Figs. 4, 5, and 6.

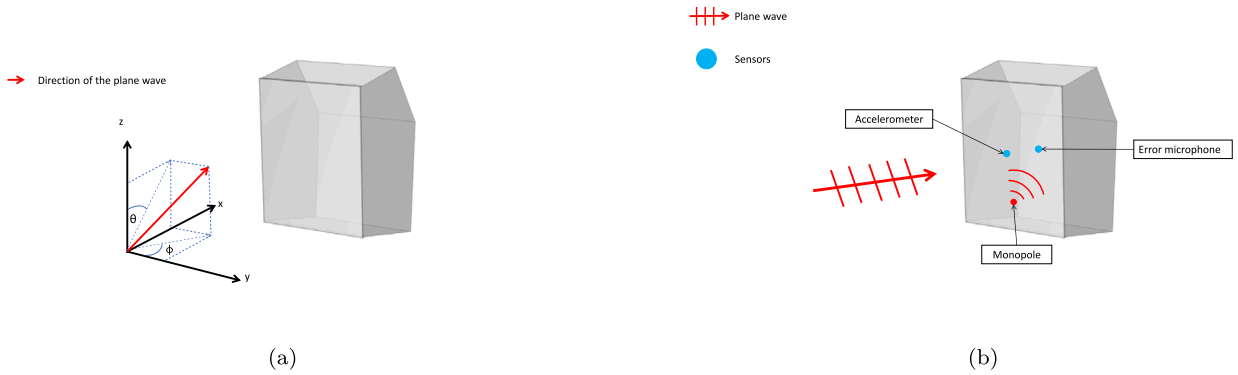


Fig. 3. Illustration of the angles θ and ϕ describing the direction of the plane wave (a). Scheme of all modeled sources: A plane wave emanates from outside the cavity to represent the disturbance, while a monopole inside the cavity simulates the canceling loudspeaker for the secondary path (b).

3.2. Identification of control paths

This final FRF computation is crucial for deriving the state-space representation, which is integral to implementing a control algorithm. To establish the transfer function, we initially estimate the frequency response of the primary, secondary, and reference paths in a residue-pole format, as indicated in the following equation.

$$H(s) = \sum_n \frac{C_n}{s - A_n} + D \tag{8}$$

where $s = i\omega f$, C_n denote the residues, A_n represent the poles, and D signifies the direct term. Our aim is to approximate these coefficients to match the previously computed frequency response behavior. To achieve this, we employ the AAA (Adaptive Antoulas-Anderson) algorithm [48]. Essentially, this algorithm furnishes a rational approximation of complex-valued data, such as our frequency response functions $H_i = H(j\omega_i)$, in a barycentric form, as shown in Eq. (9). It is noteworthy that the resulting transfer function from this method is in the continuous domain.

$$r(s) = \frac{n(s)}{d(s)} = \frac{\sum_{i=1}^{N} \frac{a_i H_i}{s - s_i}}{\sum_{i=1}^{N} \frac{a_i}{s - s_i}} \tag{9}$$

In this scenario, $N \geq 1$ represents the order of the rational function approximation. a_i signifies a series of real or complex weights associated with each support point. The terms $n(s)$ and $d(s)$ denote the partial fractions constituting the numerator and the denominator of the rational function, respectively. Here, the function $r(s)$ acts as an interpolator for the transfer function at the node $s = j\omega_i$, accommodating any non-zero weight a_i. In practical situations, supplementary data points are introduced to determine the weights a_i. These weights are optimized to minimize the linearized least squares deviation of $H(s)d(s) - n(s)$ across the additional data points. The algorithm employs a greedy strategy, systematically increasing the order by introducing an extra node $j\omega_{N+1}$ where the norm of the residual $H(s) - n(s)/d(s)$ is maximized. The solution for the weights described in Eq. (9) is achieved using a Singular Value Decomposition (SVD) technique.

$$\min \sum_{k=0}^{M} \left| \sum_{i=0}^{N} \frac{a_i \hat{H}_k}{j\hat{\omega}_k - j\omega_i} - \sum_{i=0}^{N} \frac{a_i H_i}{j\hat{\omega}_k - j\omega_i} \right|, \quad \|a\| = 1 \tag{10}$$

Here, $\hat{\omega}$ represents a vector comprising $M + 1$ frequency points distinct from the $N + 1$ frequency points ω used as nodes in the barycentric form. Similarly, $\hat{H}_k = H(j\hat{\omega}_k)$ represents the frequency responses obtained at these supplementary frequencies. The vector a containing the weights a_i is determined by solving the least squares problem depicted in Eq. (10) [49,50]. After estimating the frequency response, a state-space representation of the model is computed to facilitate its integration into a control algorithm. Initially, the state-space model is in continuous domain, described as follows:

$$\dot{x} = Ax + Bu \tag{11}$$

$$y = Cx + Du \tag{12}$$

Here, A is the state matrix, B is the input matrix, C is the output matrix, and D is the direct transfer matrix. u is the input vector, y is the output vector and x is the state vector. In order to pass from the s domain to the z domain, we used the Zero-Order Hold (ZOH) method. This method approximates the continuous-time system by assuming that the input signal is held constant between sampling intervals, thereby discretizing the system's dynamics. The ZOH method involves sampling the system's impulse response and applying a transformation to obtain the discrete-time equivalent, which allows for the analysis and design of digital control systems [51]. The conversion is done with a sampling rate of 8000 Hz.

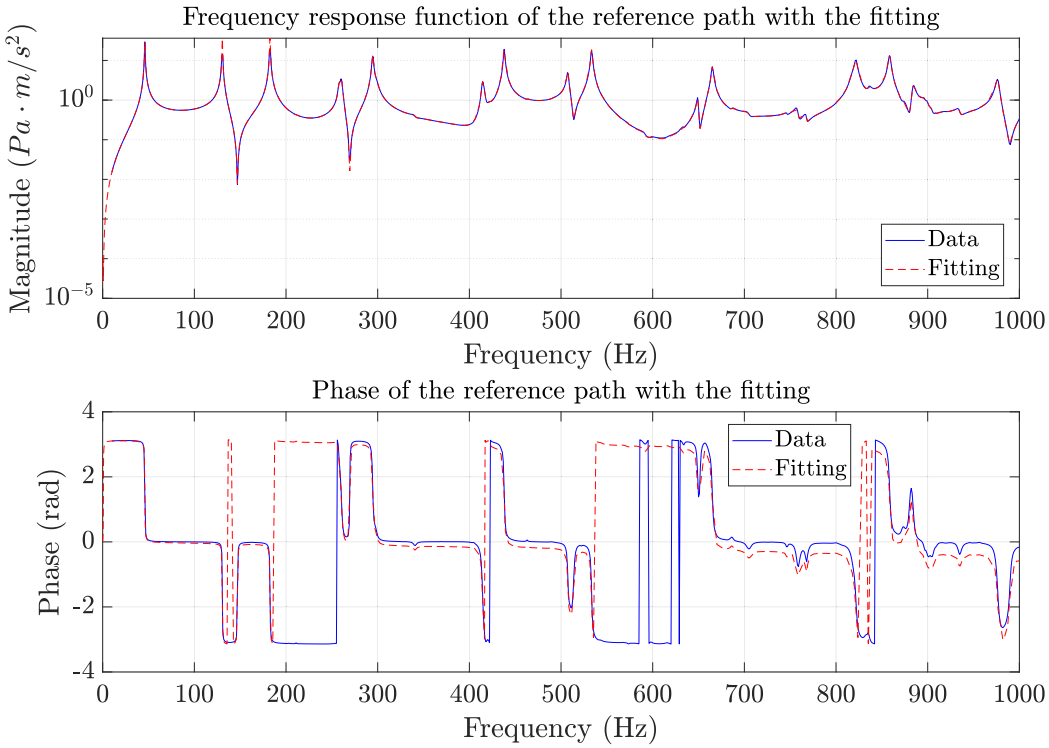


Fig. 4. Frequency response of the reference path.

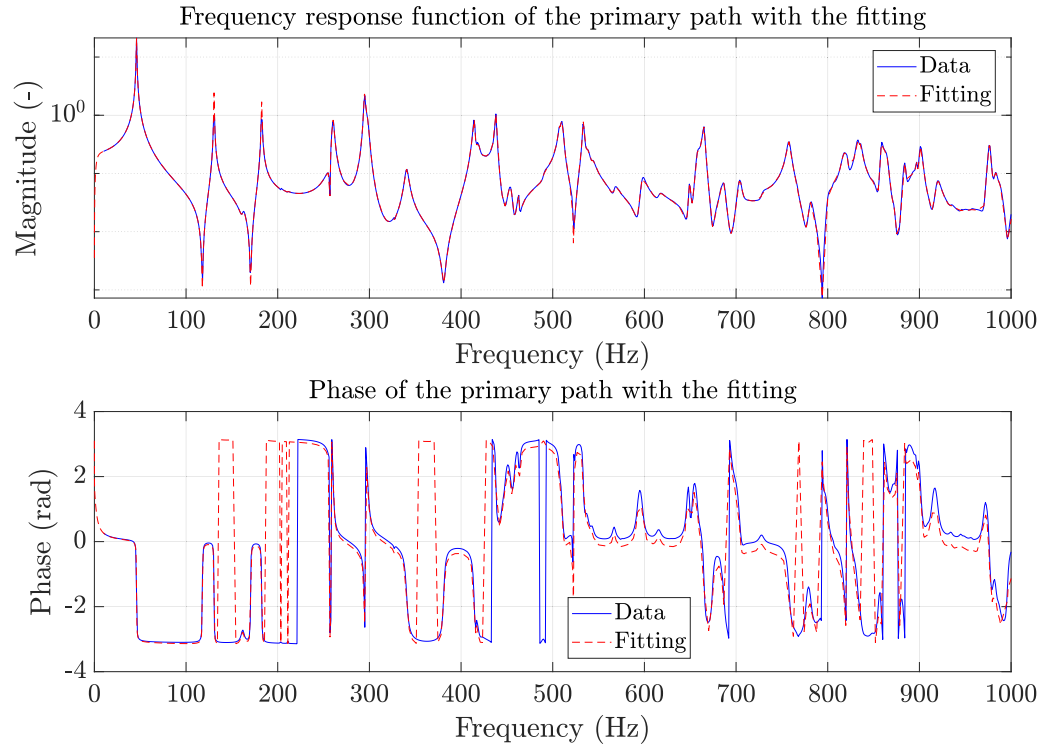


Fig. 5. Frequency response of the primary path.

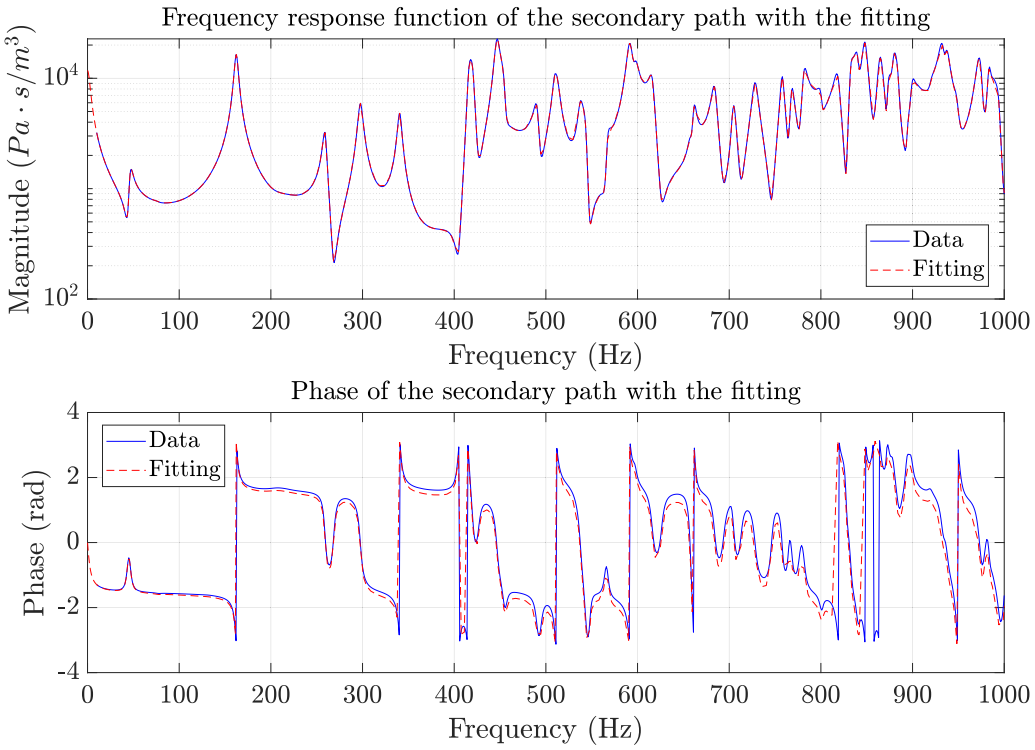


Fig. 6. Frequency response of the secondary path.

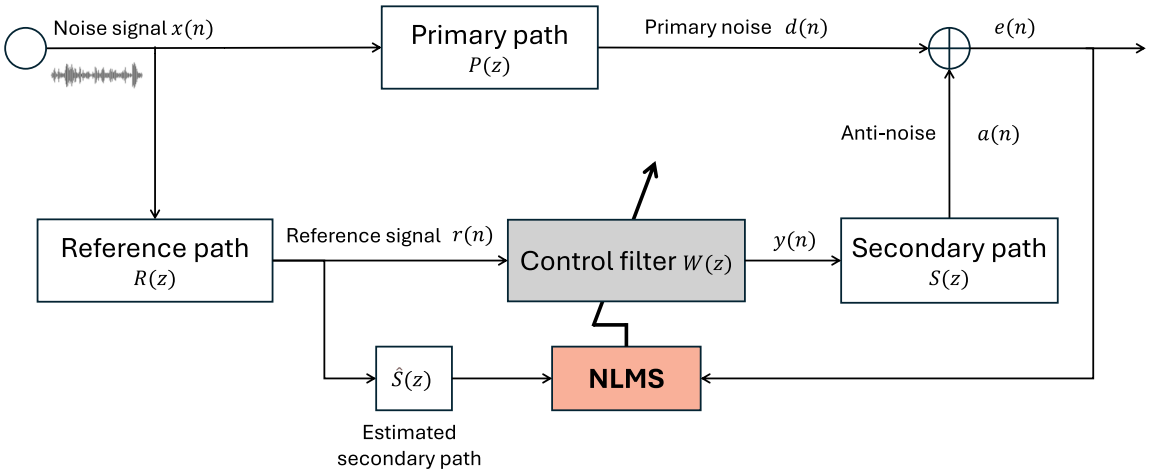


Fig. 7. FxNLMS for broadband feedforward system. The term $\hat{S}(z)$ represents the secondary path estimate, which is a transfer function that accounts for secondary path effects. $R(z)$, $P(z)$ and $S(z)$ are the transfer functions of the reference path, the primary path and the secondary path.

4. Active noise control algorithm on non-minimum phase system

4.1. FxNLMS on non-minimum phase system

The Filtered-x Normalized Least Mean Square (FxNLMS) algorithm is an adaptive algorithm proposed to compensate for the influence of the secondary path in an ANC system. This section presents the FxNLMS algorithm used in single-channel feedforward ANC applications. The block diagram of the FxNLMS algorithm is shown in Fig. 7. In the figure, $P(z)$ denotes the transfer function of the primary path, $W(z)$ represents the control filter, $S(z)$ stands for the transfer function of the secondary path, and its estimate is $\hat{S}(z)$.

The vector $x(n)$ represents the reference signal and is defined as $[x(n), x(n-1), \ldots, x(n-N+1)]$, $r(n)$ is the signal after the reference path. Let T denote the transpose operation and N denotes the length of the control filter. The resulting output of the control filter is therefore $y(n)$ and is defined in Eq. (13).

$$y(n) = w^T(n)r(n) \tag{13}$$

where $w(n)$ represents the coefficient vector of the control filter. The residual error signal can be expressed as follows:

$$e(n) = d(n) - s(n) * \left(w^T(n)r(n)\right) \tag{14}$$

where $d(n)$ and $s(n)$ denote the primary disturbance and the impulse response of the secondary path, respectively, and $*$ is the linear convolution. To minimize the instantaneous squared error, $J(n) = e^2(n)$, the most widely used method to achieve this objective is the gradient descent method, which updates the coefficient vector in the negative gradient direction with step size μ:

$$w(n+1) = w(n) - \frac{\mu}{2}\nabla J(n) \tag{15}$$

where $\nabla J(n)$ denotes the instantaneous estimate of the MSE gradient at time n and can be expressed as follows:

$$\nabla J(n) = \nabla e^2(n) = -2e(n)r'(n) \tag{16}$$

The filtered reference signal vector is given by:

$$r'(n) = \hat{s}(n) * r(n) \tag{17}$$

where the impulse response $s(n)$ is replaced with its estimate $\hat{s}(n)$. Hence, substituting Eqs. (16) and (17) into (15) yields the update equation of the FxNLMS algorithm as follows:

$$w(n+1) = w(n) + \mu e(n)\frac{r'^*(n)}{r'^H(n)r'(n) + \epsilon} \tag{18}$$

Here, $*$ denotes the complex conjugate and H represents the Hermitian transpose. ϵ is a small positive constant that overcomes the possible numerical instability when updating the weights. It prevents the denominator from approaching zero. In our case, it is equal to 2.2204×10^{-16}. If the group delay of the secondary path is assumed to be D_s, the bound of the step size μ in the FxNLMS algorithm is given by:

$$0 < \mu < \frac{1}{\lambda_{\max} D_s} \tag{19}$$

where $\lambda_{\max}$ represents the maximum eigenvalue of the auto-correlation matrix $R_{r'} = E[r'^T(n)r'(n)]$ of the filtered reference signal [52].

The optimal controller for ANC is ideally given by $W(z) = \frac{P(z)}{S(z)}$. However, when inverting the secondary path $S(z)$ in non-minimum phase systems, unstable zeros outside the unit circle become unstable poles, resulting in phase distortions that complicate the estimation of the secondary path $\hat{S}(z)$, which is crucial for FxNLMS. This introduces poles into the feedback loop, which can lead to instabilities in the adaptive filter and complicate weight management, reducing ANC performance. In addition, these systems have different group delays that further distort the filtered reference signal $r'(n)$ as defined in Eq. (17). These distortions can impair the adaptation of the control filter $W(z)$, leading to potential instability and divergence. As a result, undershoot and delays may occur during control application, as the algorithm struggles to address the inherent phase and amplitude distortions, as shown in Fig. 8.

4.2. Kalman filter

To overcome the limitations of the digital NLMS filter in non-minimum phase systems, a Kalman filter was chosen as an alternative. The internal model technique uses secondary path estimation to recover the disturbance from the error signal, which allows the use of the traditional normalized least mean squares (NLMS) algorithm for noise control. This modified configuration can also be used to implement the Kalman filter method, replacing the NLMS model in Fig. 7 with the Kalman filter to complete the control filter update.

To implement the Kalman filter method, the state function must first be defined. The following sections describe the definition of the state function for the ANC system, which is adapted to the problem under consideration. The authors in [53] give a more complete definition of the Kalman filter. As known, when the adaptive algorithm converges, the ANC system should reach the optimal control filter, which is a constant solution given by:

$$w(n+1) = w(n) = w_o, \tag{20}$$

where w_o denotes the optimal control filter. Meanwhile, the attenuated noise can be expressed as

$$e_o(n) = d(n) - \sum_{i=0}^{L-1} \hat{s}_i r^T(n-i)w_o(n), \tag{21}$$

$e_o(n)$ is the difference between the primary noise and the anti-noise. Let $d(n)$ and $r(n)$ denote the disturbance signal and the signal after the reference path, respectively, while $\hat{s}$ represents the i th coefficient of the secondary path estimate. Since the optimal control filter is a constant vector, (21) can be rewritten as

$$d(n) = r'^T(n)w_o(n) + e_o(n), \tag{22}$$

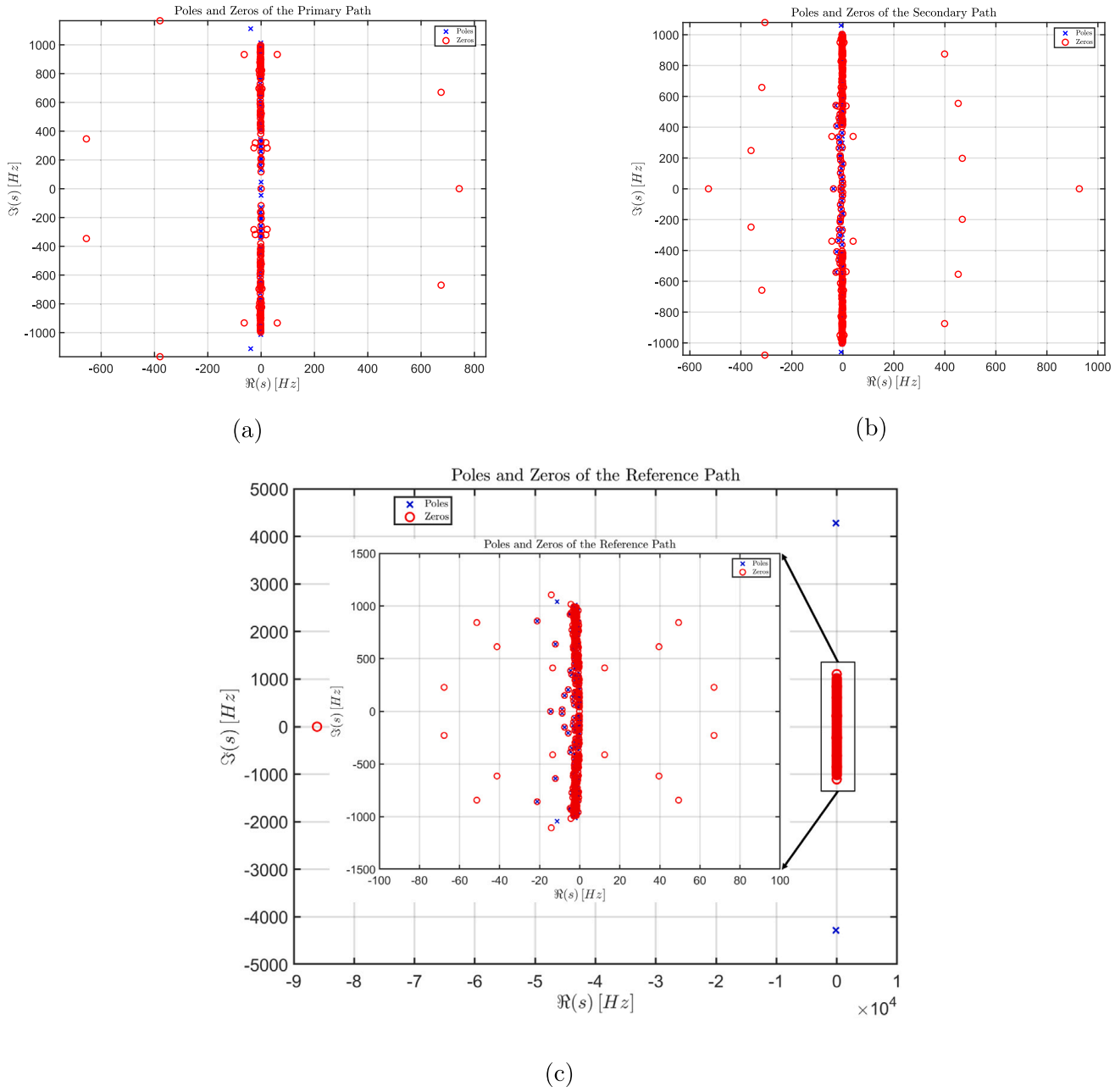


Fig. 8. Pole-zero diagrams in the s-domain for the primary path (a), secondary path (b), and reference path (c).

and the filtered reference signal is given by

$$r'(n) = \sum_{i=0}^{L-1} \hat{s}_i r(n-i). \tag{23}$$

It is obvious to use (20) and (22) as state equation and observation equation, respectively. The recursive equations of the Kalman filter are as follows:

(i) The prediction of the new state is given by:

$$\hat{w}(n) = w(n-1) \in \mathbb{R}^{N\times 1}, \tag{24}$$

where $\hat{w}(n)$ represents the predicted control filter vector at time n. It is an estimate made before incorporating the new observation, relying solely on the previous state. This prediction assumes that the system's dynamics remain unchanged between time steps. The term $w(n-1)$ corresponds to the estimated control filter vector from the previous time step $n-1$. N denotes the length of the control filter.

Table 2
Kalman filter algorithm.

Step	Equation
Prediction of the new state	$\hat{w}(n) = w(n-1) \in \mathbb{R}^{N\times 1}$
Prediction of the state-error auto-correlation matrix	$P_{\text{pred}}(n) = P(n-1) \in \mathbb{R}^{N\times N}$
Kalman gain matrix	$K(n) = P_{\text{pred}}(n)r'(n)\left[r'(n)^T P_{\text{pred}}(n)r'(n) + q(n)\right]^{-1}$
State estimate	$\hat{w}(n+1) = \hat{w}(n) + K(n)\left[d(n) - r'(n)^T \hat{w}(n)\right]$
Update of the state-error auto-correlation matrix	$P(n) = \left(I - K(n)r'(n)^T\right) P_{\text{pred}}(n)$

(ii) The prediction of the auto-correlation matrix of the state error is given by:

$$P_{\text{pred}}(n) = P(n-1) \in \mathbb{R}^{N\times N}, \tag{25}$$

where $P_{\text{pred}}(n)$ is the predicted auto-correlation matrix of the state error at time n. $P(n-1)$ is the auto-correlation matrix of the state error at time $n-1$, which is used to predict the value at the current time step. Here, it is assumed that the variance of the final state error equals zero. In order to fully specify the Kalman filter recursion, the state-error covariance matrix $P(n-1)$ must be initialized. A typical approach, which reflects a lack of certainty about the state, is to set the state-error covariance matrix as a scaled identity matrix $P(n-1) = \delta \cdot I$, where δ is a positive constant and I denotes the identity matrix of dimensions $N \times N$.

(iii) The Kalman gain matrix is calculated as:

$$K(n) = P_{\text{pred}}(n)r'(n)\left[r'(n)^T P_{\text{pred}}(n)r'(n) + q(n)\right]^{-1}, \tag{26}$$

where $K(n)$ is the Kalman gain matrix at time n and $q(n)$ represents the variance of the observation error, defined as:

$$q(n) = E[e_o^2(n)]. \tag{27}$$

(iv) The state estimate is given by:

$$\hat{w}(n+1) = \hat{w}(n) + K(n)\left[d(n) - r'(n)^T \hat{w}(n)\right] \tag{28}$$

where $w(n)$ is the updated state estimate at time n, which incorporates the Kalman gain and measurement correction. $\hat{w}(n)$ is the prior state estimate at time n.

(v) The auto-correlation matrix of the state error is updated as:

$$P(n) = \left(I - K(n)r'(n)^T\right) P_{\text{pred}}(n) \tag{29}$$

where $P(n)$ is the updated auto-correlation matrix of the state error at time n, which reflects the uncertainty in the updated state estimate. The term $I - K(n)r'(n)^T$ represents the correction to the prior prediction based on the measurement update, where I is the identity matrix. This update reduces the uncertainty in the state estimate after considering the measurement information. The steps of the Kalman filter algorithm are summarized in Table 2.

The Kalman filter algorithm can outperform the FxNLMS algorithm due to its recursive and adaptive nature, which allows it to handle the uncertainties in the system more effectively. Unlike FxNLMS, which adjusts the filter based on a fixed secondary path model and a gradient descent update, the Kalman filter continuously refines both the state estimate and its associated uncertainty. This is achieved through the use of the Kalman gain matrix $K(n)$, calculated in Eq. (26), which optimally weighs the new measurement relative to the prior estimate, allowing the algorithm to adjust to changing conditions in the secondary path. Furthermore, the state prediction step, given by Eq. (24), allows the Kalman filter to incorporate prior knowledge of the system dynamics, leading to more accurate predictions of the control filter state. Updating the autocorrelation matrix of the state error, described by Eq. (29), allows the Kalman filter to dynamically adjust the estimation uncertainty based on new measurements, effectively reducing the impact of modeling errors or perturbations on the secondary path. This mechanism, which is not present in FxNLMS, allows the Kalman filter to better mitigate non-minimum phase behavior and disturbances, ultimately improving the noise rejection performance and overall accuracy of the system. While the Kalman filter shows strong performance under modeling uncertainties, other approaches also address variations in the secondary path. These include modeless algorithms, which bypass explicit path modeling using feedback structures, and online modeling techniques that continuously update the path estimate. For example, in [54], the authors use fractional signal processing for online adaptation, and [55] propose a simultaneous equation-based method for real-time modeling.

5. Results and discussion

To evaluate and compare the performance of the FxNLMS and Kalman filter algorithms, the results are presented in terms of noise reduction achieved with broadband noise (20–450 Hz). Additionally, the influence of the noise's frequency range on control performance was assessed by applying both algorithms to narrowband noise in low, mid, and high-frequency ranges. Presenting these results will highlight the specific characteristics of each algorithm and the control environment.

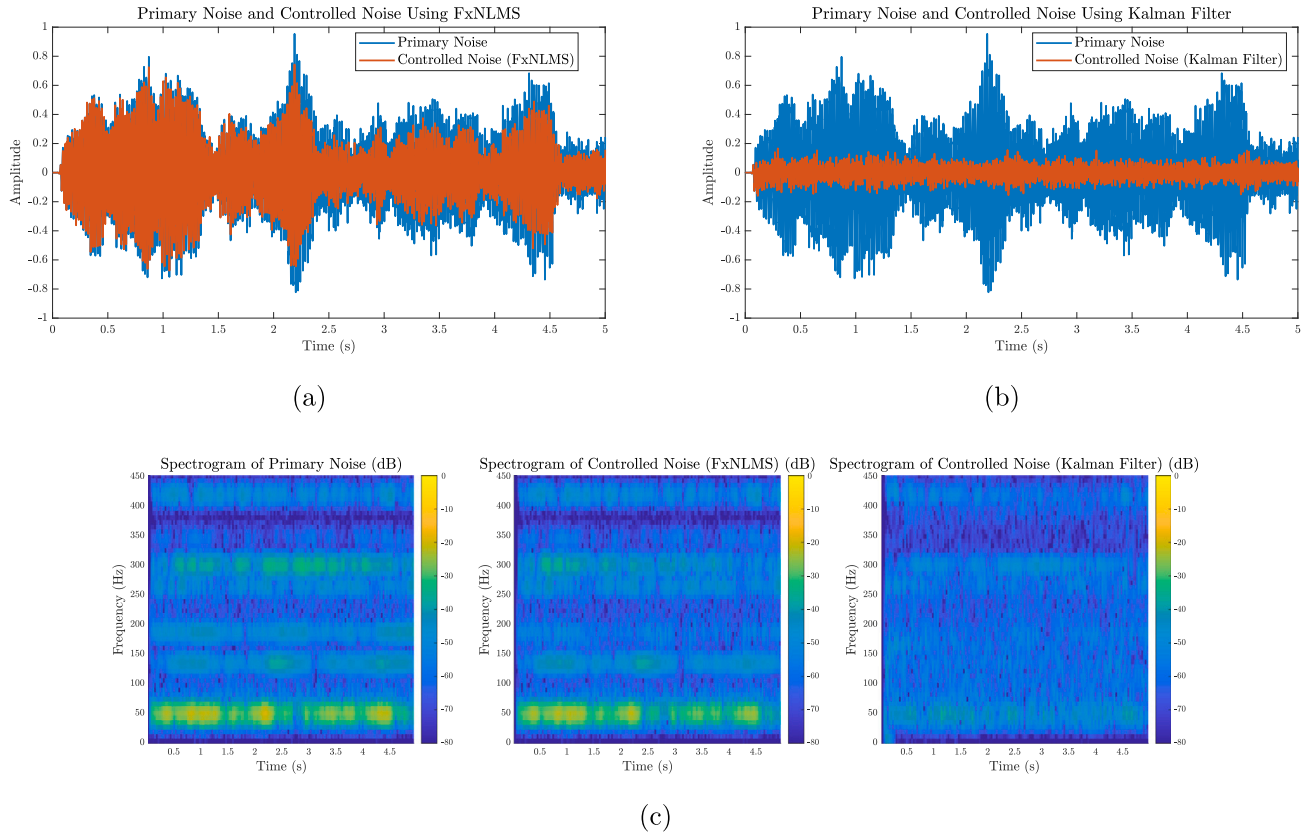


Fig. 9. Graphical representation of the control. (a) and (b) show the time domain signals of the results with FxNLMS and with the Kalman filter, respectively. (c) shows the spectrograms of the primary noise, the controlled noise using FxNLMS, and the controlled noise using the Kalman filter.

5.1. Noise reduction comparison

Concerning the results obtained with both algorithms in our non-minimum phase system, the control was applied using a disturbance defined as a filtered random noise. The noise is Gaussian with a mean of zero and a unit variance. It is filtered by a low-pass filter up to 450 Hz, with a passband edge frequency of 400 Hz and a stopband frequency of 450 Hz. The maximum passband ripple and the minimum stopband attenuation are set at 0.1 dB and 80 dB, respectively. All simulations are performed with the sampling rate of the state space, which is 8 kHz. The FxNLMS algorithm uses a step size μ of 0.00042 and a filter length of 512. The Kalman filter has a variance, as referenced in Eq. (27), of 0.5. Simulations were conducted on signals of 5 s in length. Fig. 9 presents the results obtained under these conditions.

As presented in Fig. 9, the use of the Kalman filter is more suitable for our system. The spectrograms show that the control is consistent across the frequency range. Indeed, the mean power of the controlled noise using the Kalman filter is significantly lower than that of the FxNLMS, with −63.25 dB for the Kalman filter compared to −49.80 dB for the FxNLMS. Fig. 10 compares the Power Spectral Density (PSD) of the primary noise, the controlled noise using FxNLMS, and the controlled noise using the Kalman filter. When plotting the PSD, we observe a waterbed effect, which is typical of non-minimum phase systems. The waterbed effect refers to the phenomenon where reducing the power in one frequency band results in an increase in power at another frequency band, much like pressing down on one area of a waterbed causes another area to rise. This effect is evident in our system, as the Kalman filter effectively controls the noise amplified by the resonance frequency of the primary path, as shown in the Frequency Response Function (FRF) in Fig. 5. However, at the anti-resonance frequencies, the primary noise can be lower than the controlled noise. While this configuration offers convenient control, it also means that when the noise is below a certain threshold, it will not be effectively controlled. Given the unique characteristics of our system, reducing the observation variance would lead to an undershoot at the beginning of the control. Fig. 10 also compares the Mean Square Error (MSE) of the noise reduction achieved using the FxNLMS algorithm and the Kalman filter. As illustrated in the figure, the noise control performance is significantly better with the Kalman filter. The improvement in noise reduction can be as much as 15.26 dB.

5.2. Influence of the frequency range

To evaluate the influence of the disturbance's frequency range, noise control was applied using both algorithms across different frequency bands. The frequency ranges considered in this study are 0–450 Hz, 0–225 Hz, 225–450 Hz, 0–150 Hz, 150–300 Hz, and 300–450 Hz. This approach allows us to observe how the system reacts to different components of the disturbance across specific

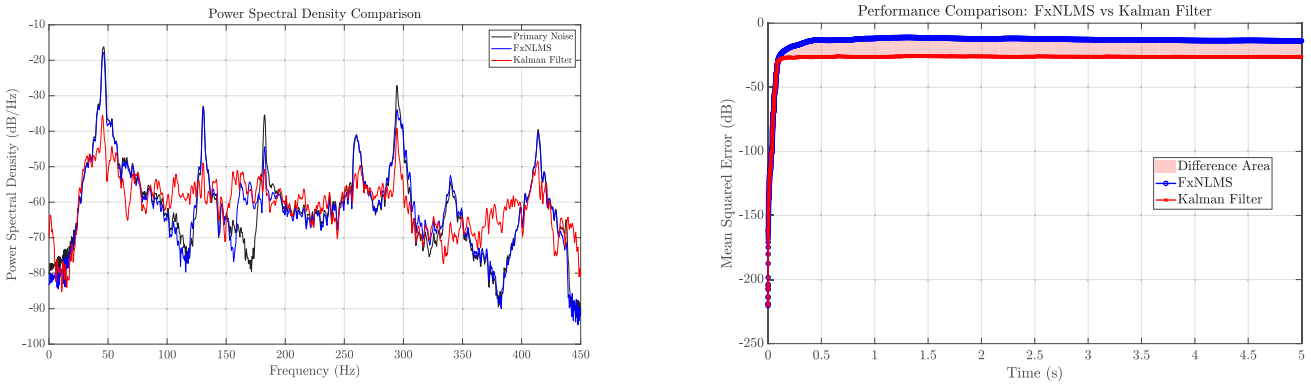


Fig. 10. Comparison of noise reduction performance between FxNLMS and the Kalman filter. (a) Power spectral density (PSD) representation of the primary noise and the controlled noise using FxNLMS and the Kalman filter. (b) Mean square error (MSE) of the controlled noise using FxNLMS and the Kalman filter.

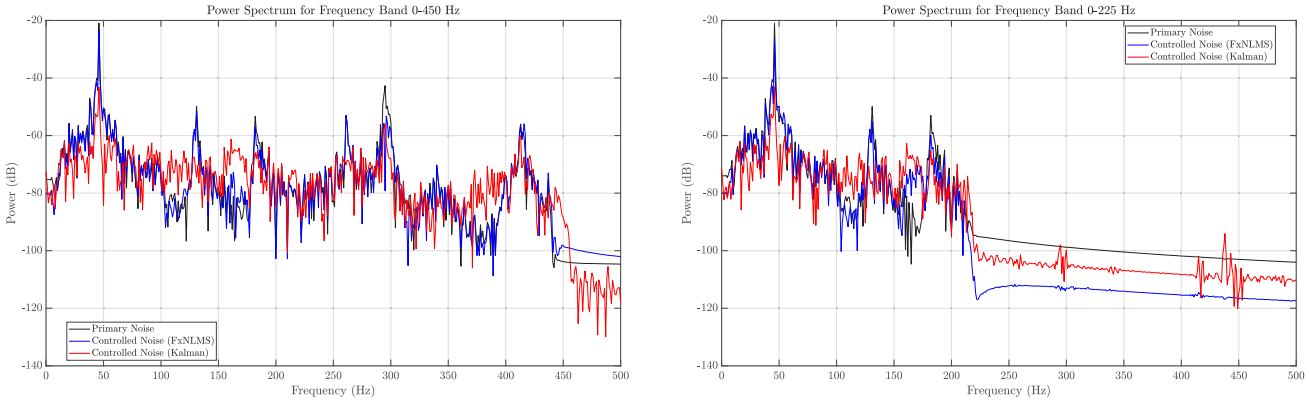


Fig. 11. Power Spectra for Frequency Bands 0–450 Hz and 0–225 Hz.

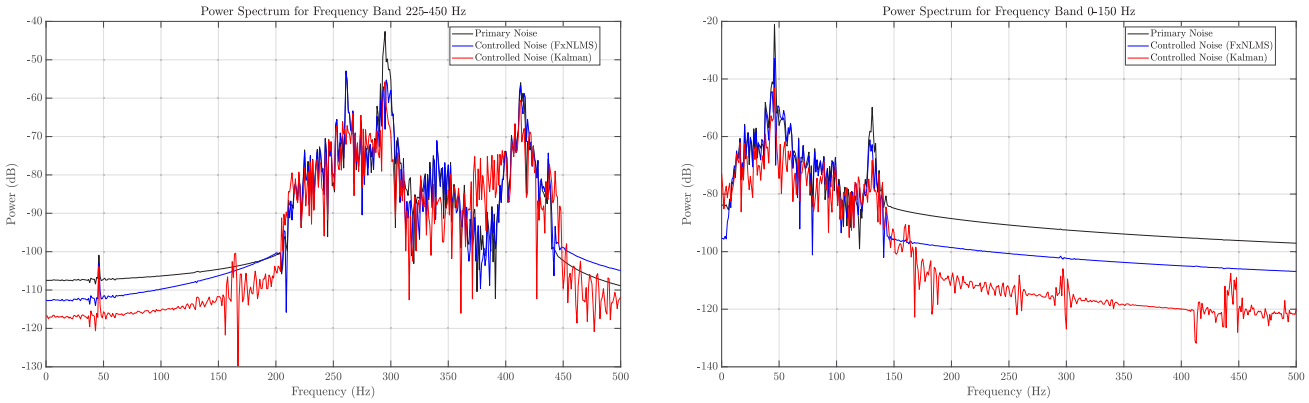
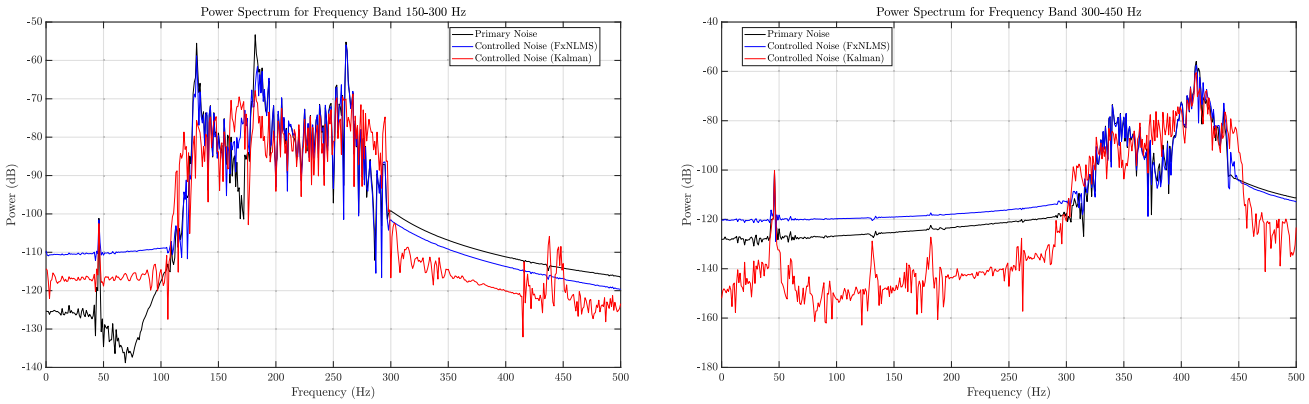


Fig. 12. Power Spectra for Frequency Bands 225–450 Hz and 0–150 Hz.

frequency ranges. The results of the noise control performance are illustrated in Figs. 11, 12, and 13, which show the power spectra corresponding to the different frequency bands. These figures allow a detailed comparison of the control effectiveness for each frequency range, highlighting the influence of the disturbance's frequency content on the system's performance.

In these figures, the waterbed effect, characteristic of non-minimum phase systems, is visible. The signal controlled with the Kalman filter (shown in red in the figure) exhibits a lower noise level at resonance frequencies and a higher level at anti-resonance



(a) Power Spectrum for Frequency Band 150-300 Hz

(b) Power Spectrum for Frequency Band 300-450 Hz

Fig. 13. Power Spectra for Frequency Bands 150–300 Hz and 300–450 Hz.

Table 3
Comparison of SPL and Noise Reduction for different Frequency Ranges.

Frequency Range (Hz)	SPL Primary Noise (dB)	SPL FxNLMS (dB)	SPL Kalman (dB)	NR FxNLMS (dB)	NR Kalman (dB)
0–450	80.39	77.67	62.47	2.72	16.92
0–225	80.31	74.67	62.07	5.64	18.24
225–450	65.34	60.57	55.65	4.77	9.69
0–150	80.26	71.55	61.25	8.71	19.01
150–300	60.80	56.59	50.72	4.16	10.03
300–450	53.88	53.34	49.68	0.53	4.19

frequencies, resulting in a power spectrum that tends to flatten. Additionally, it is evident that noise control is more effective in the lower frequency range compared to the higher frequency range. This is due to the higher presence of modes at high frequencies compared to low frequencies.

To highlight this phenomenon, the Sound Pressure Level (SPL) for each frequency range was calculated, compared to the controlled signals from each algorithm, and the resulting Noise Reduction (NR) was deduced. The results are shown in Table 3 and Eq. (30) was used to calculate the SPL.

$$\text{SPL (dB)} = 20\log_{10}\left(\frac{p}{p_0}\right) \tag{30}$$

where p is the root-mean-square sound pressure, and p_0 is the reference sound pressure (typically 20 μPa in air).

As shown in the table, the Kalman filter is more effective than FxNLMS in all cases studied. It is noteworthy that noise reduction is lower for both filters at higher frequencies. This is particularly evident when the noise is filtered between 225–450 Hz and 300–450 Hz, where the noise reductions achieved with FxNLMS are 4.77 dB and 0.53 dB, respectively. With the Kalman filter, the reductions are 9.69 dB and 4.19 dB. In the 300–450 Hz frequency range, the noise reduction is particularly poor. This result is due to the undershoot systematically observed with FxNLMS in this frequency range. Although undershoots can also occur with the Kalman filter, their amplitude is significantly lower than those observed with FxNLMS. Fig. 14 shows the bar chart of the NMSE values calculated using Eq. (31).

$$\text{NMSE} = 10\log_{10}\left(\frac{\sum_{t=1}^{L} e^2(t)}{\sum_{t=1}^{L} d^2(t)}\right) \tag{31}$$

It illustrates the control efficiency achieved across different frequency ranges. Both algorithms, FxNLMS and the Kalman filter, demonstrate less effective performance at higher frequencies, as indicated by the increasing NMSE values in those bands.

5.3. Computational complexity

In this section, we present a comparative analysis of the computational complexity and noise reduction performance of three algorithms: FxNLMS, FxRLS, and the Kalman filter. While the performance of the FxNLMS algorithm is insufficient for more demanding applications, the FxRLS delivers promising results, and the Kalman filter demonstrates strong performance, particularly in more challenging scenarios such as NMP systems. The FxRLS algorithm, introduced in [56], requires $4N^2 + 6N$ multiplications and $3N^2 + 3N$ ddditions, where N is the number of coefficients in the control filter. The FxNLMS algorithm demands $4N + 1$ multiplications and $4N + 3$ additions. While FxNLMS offers a computationally efficient solution, its noise reduction performance, at 2.30 dB, does not meet the needs of more demanding applications. The FxRLS algorithm, on the other hand, strikes a better balance between performance and computational efficiency, with a noise reduction of 10.25 dB. This makes it a more viable option for

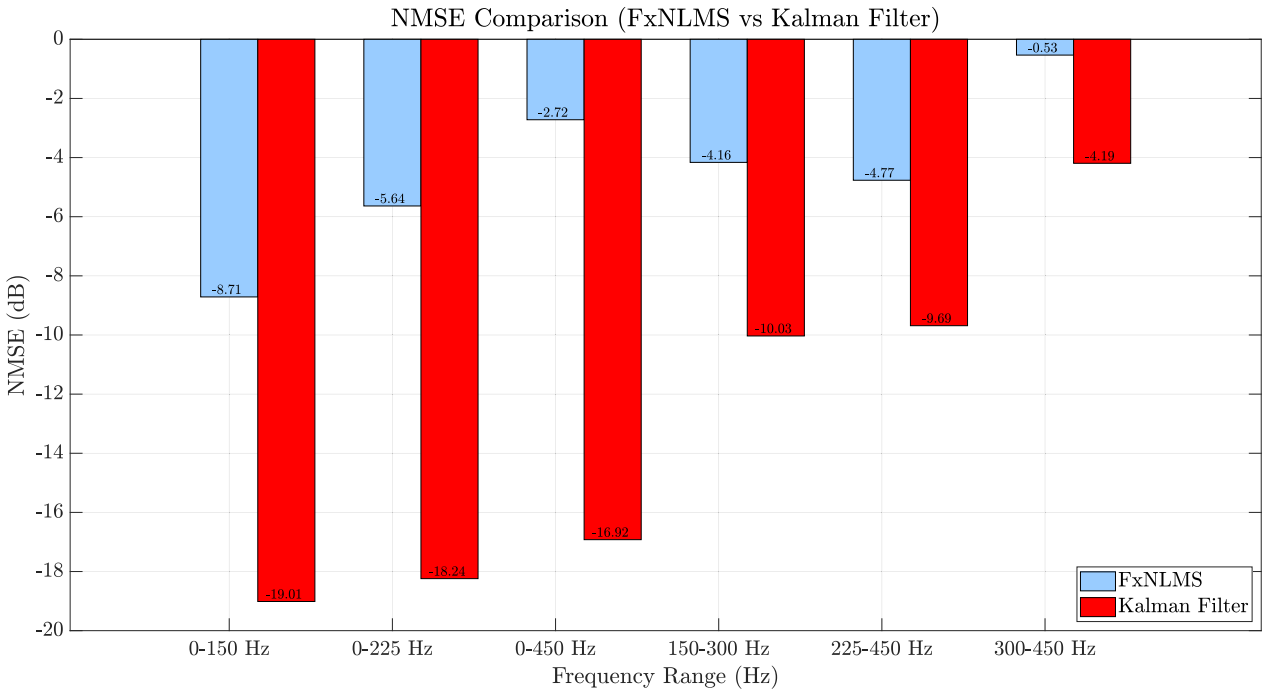


Fig. 14. Bar chart of the NMSE obtained for different frequency ranges using FxNLMS and the Kalman filter.

Table 4
Complexity analysis of FxNLMS, FxRLS, and Kalman filter algorithms.

Algorithm	FLOPs Multiplication	FLOPs Addition	Noise Reduction (dB)
FxNLMS	$5N+1$	$5N+3$	2.30
FxRLS	$7N^2+6N$	$4N^2+8N$	10.25
Kalman Filter	N^3+2N^2+4N+1	N^3+3N^2+3N-1	17.10

scenarios where higher performance is required without a prohibitive computational cost. In contrast, the Kalman filter, although computationally more demanding, provides strong noise reduction performance. The Kalman filter requires N^3+2N^2+4N+1 multiplications and N^3+2N^2+3N-1 additions. Despite its higher computational cost, the Kalman filter achieves a noise reduction of 17.10 dB, outperforming both the FxRLS and FxNLMS. This makes the Kalman filter a suitable choice when substantial noise reduction is necessary. The RLS filter in this study was implemented with an exponential weighting factor of $\lambda = 0.9999$, and the control was applied to random noise in the frequency range between 50 Hz and 450 Hz. This comparison highlights the trade-off between computational cost and performance, emphasizing the Kalman filter's effectiveness in noise reduction, especially in more challenging environments (see Table 4).

6. Conclusion

The paper presents a FEM model of an encapsulated structure referred to as the Noise Box. The model is designed to simulate the interior of a car and serves as the environment where noise control is applied. A key challenge of this environment is that the system model exhibits non-minimum phase characteristics, which do not directly affect the system's stability but make noise control more difficult due to its complex dynamic behavior. FxNLMS algorithm result in poor noise reduction, which motivates the use of a Kalman filter to achieve better efficiency.

However, the method has some limitations. Even with the Kalman filter, undershoots and reduced noise suppression effectiveness may occur at higher frequencies. This is mainly due to the fact that noise at higher frequencies is more difficult to control due to the shorter wavelengths and more complex interactions of sound waves in an enclosed space. In addition, the non-minimum phase of the system leads to delays and phase shifts that make precise control at these frequencies more difficult. Therefore, while the Kalman filter shows better performance than conventional methods, reduction in higher frequency bands remains a persistent challenge, requiring further optimization or the combination of multiple control strategies. Another important consideration is the computational complexity of the Kalman filter.

In future studies, more robust algorithms capable of controlling more complex noise will be employed. The use of artificial intelligence in active noise control could be a viable option due to its ability to capture data patterns. Additionally, a hybrid solution will be explored, integrating artificial intelligence with traditional algorithms. Future research should also focus on optimizing its implementation, to reduce computational load while preserving its noise reduction capabilities.

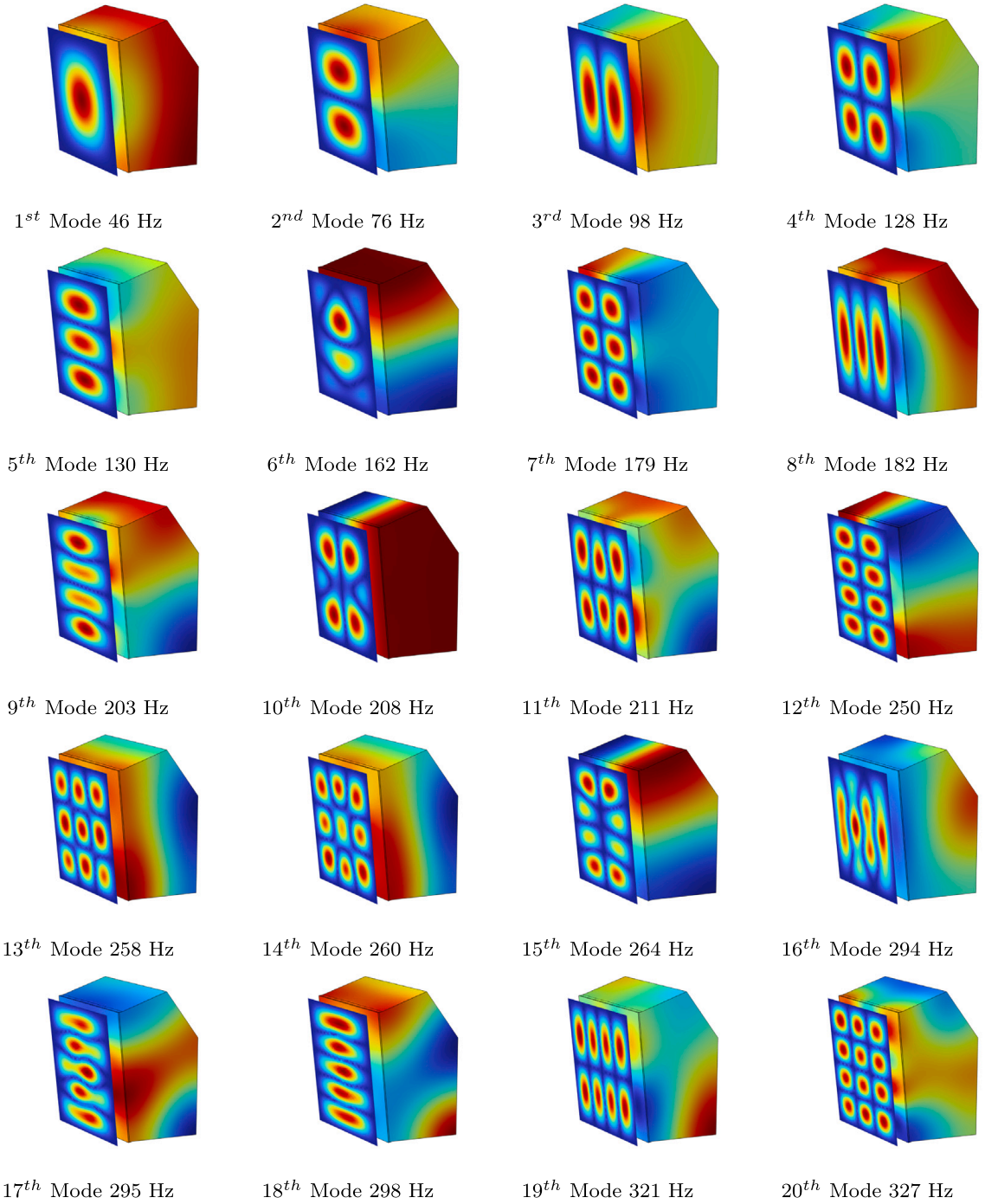


Fig. A.15. Noise Box mode shapes.

CRediT authorship contribution statement

Alkahf Aboutiman: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Reza Shams:** Writing – review & editing, Validation, Resources, Formal analysis, Data curation, Conceptualization. **Hamid Reza Karimi:** Writing – review & editing, Writing – original draft, Validation, Supervision, Conceptualization. **Francesco Ripamonti:** Writing – review & editing, Writing – original draft, Validation, Supervision, Conceptualization. **Marek Pawełczyk:** Writing – review & editing, Writing – original draft, Validation, Supervision, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The work was supported by the project IN-NOVA: Active reduction of noise transmitted into and from enclosures through encapsulated structures, which has received funding from the European Union's Horizon Europe programme under the Marie Sklodowska-Curie grant agreement no. 101073037.

Annex. Noise box mode shapes

See Fig. A.15.

Data availability

Data will be made available on request.

References

[1] R.H. Cabell, C. Fuller, A principal component algorithm for feedforward active noise and vibration control, J. Sound Vib. 227 (1) (1999) 159–181.
[2] G.C. Goodwin, E.I. Silva, D.E. Quevedo, Analysis and design of networked control systems using the additive noise model methodology, Asian J. Control 12 (4) (2010) 443–459.
[3] W.M. Hartmann, Signals, Sound, and Sensation, Springer Science & Business Media, 2004.
[4] Y. Li, X. Wang, R. Huang, Z. Qiu, Active vibration and noise control of vibro-acoustic system by using PID controller, J. Sound Vib. 348 (2015) 57–70.
[5] J.F. Wilby, Aircraft interior noise, J. Sound Vib. 190 (3) (1996) 545–564.
[6] C. Fuller, C. Hansen, S. Snyder, Active control of sound radiation from a vibrating rectangular panel by sound sources and vibration inputs: an experimental comparison, J. Sound Vib. 145 (2) (1991) 195–215.
[7] A. Gonzalez, M. Ferrer, M. De Diego, G. Pinero, J. Garcia-Bonito, Sound quality of low-frequency and car engine noises after active noise control, J. Sound Vib. 265 (3) (2003) 663–679.
[8] M. Misol, S. Algermissen, H.P. Monner, Experimental investigation of different active noise control concepts applied to a passenger car equipped with an active windshield, J. Sound Vib. 331 (10) (2012) 2209–2219.
[9] Y. Li, L. Cheng, Active noise control of a mechanically linked double panel system coupled with an acoustic enclosure, J. Sound Vib. 297 (3–5) (2006) 1068–1074.
[10] P. Shorter, Recent advances in automotive interior noise prediction, SAE BRASIL Noise Vib. Conf. (2008).
[11] E. Rustighi, S. Elliott, Stochastic road excitation and control feasibility in a 2D linear tyre model, J. Sound Vib. 300 (3–5) (2007) 490–501.
[12] W. Belgacem, A. Berry, P. Masson, Active vibration control on a quarter-car for cancellation of road noise disturbance, J. Sound Vib. 331 (14) (2012) 3240–3254.
[13] T.J. Sutton, S.J. Elliott, A.M. McDonald, T.J. Saunders, Active control of road noise inside vehicles, Noise Control Eng. J. 42 (4) (1994) 137–147.
[14] Y. Wang, T. Feng, X. Wang, H. Guo, H. Qi, An improved LMS algorithm for active sound-quality control of vehicle interior noise based on auditory masking effect, Mech. Syst. Signal Process. 108 (2018) 292–303.
[15] H. Sano, T. Inoue, A. Takahashi, K. Terai, Y. Nakamura, Active control system for low-frequency road noise combined with an audio system, IEEE Trans. Speech Audio Process. 9 (7) (2001) 755–763.
[16] W. Jung, S.J. Elliott, J. Cheer, Local active control of road noise inside a vehicle, Mech. Syst. Signal Process. 121 (2019) 144–157.
[17] H. Hariri, P. Masson, L. Frechette, Road noise reduction inside car cabin using active structural acoustic control (ASAC), in: ACTUATOR; International Conference and Exhibition on New Actuator Systems and Applications 2021, VDE, 2021, pp. 1–4.
[18] J. Cheer, S.J. Elliott, The design and performance of feedback controllers for the attenuation of road noise in vehicles, Int. J. Acoust. Vib. 19 (3) (2014) 155–164.
[19] J. Cheer, S.J. Elliott, Multichannel control systems for the attenuation of interior road noise in vehicles, Mech. Syst. Signal Process. 60 (2015) 753–769.
[20] P. Ramos, A. Salinas, A. López, E. Masgrau, Practical implementation of a multiple-channel FxLMS active noise control system with shaping of the residual noise inside a van, in: Proc. ACTIVE, vol. 2, 2002, pp. 303–314.
[21] G. Sun, J. Duan, M. Li, T.C. Lim, Numerical modeling and active noise control of impact road noise inside a vehicle compartment, Noise Control Eng. J. 63 (1) (2015) 72–85.
[22] F. Ihlenburg, Sound in vibrating cabins: Physical effects, mathematical formulation, computational simulation with fem, Comput. Asp. Struct. Acoust. Vib. (2009) 103–170.
[23] D. Stanef, C. Hansen, R. Morgans, Active control analysis of mining vehicle cabin noise using finite element modelling, J. Sound Vib. 277 (1–2) (2004) 277–297.
[24] J. Sousa, C.A. Silva, J.S. da Costa, Fuzzy active noise modeling and control, Internat. J. Approx. Reason. 33 (1) (2003) 51–70.
[25] P. Sas, C. Bao, F. Augusztinovicz, W. Desmet, Active control of sound transmission through a double panel partition, J. Sound Vib. 180 (4) (1995) 609–625.
[26] J. Loncaric, V. Ryaben'kii, S. Tsynkov, Active shielding and control of noise, SIAM J. Appl. Math. 62 (2) (2001) 563–596.
[27] S.-P. Moon, T.-G. Chang, The derivation of the stability bound of the feedback ANC system that has an error in the estimated secondary path model, Appl. Sci. 8 (2) (2018) 210.
[28] C.H. Hansen, S.D. Snyder, X. Qiu, L.A. Brooks, D.J. Moreau, Active Control of Noise and Vibration, E & Fn Spon London, 1997.
[29] T. Yang, L. Zhu, X. Li, L. Pang, An online secondary path modeling method with regularized step size and self-tuning power scheduling, J. Acoust. Soc. Am. 143 (2) (2018) 1076–1084.
[30] P.A. Lopes, M.S. Piedade, A Kalman filter approach to active noise control, in: 2000 10th European Signal Processing Conference, IEEE, 2000, pp. 1–4.
[31] S. van Ophem, A.P. Berkhoff, Multi-channel Kalman filters for active noise control, J. Acoust. Soc. Am. 133 (4) (2013) 2105–2115.

[32] J. Fabry, S. Liebich, P. Vary, P. Jax, Active noise control with reduced-complexity Kalman filter, in: 2018 16th International Workshop on Acoustic Signal Enhancement, IWAENC, IEEE, 2018, pp. 166–170.
[33] L. Wu, X. Qiu, Y. Guo, A generalized leaky FxLMS algorithm for tuning the waterbed effect of feedback active noise control systems, Mech. Syst. Signal Process. 106 (2018) 13–23.
[34] A.A. Milani, G. Kannan, I.M. Panahi, On maximum achievable noise reduction in ANC systems, in: 2010 IEEE International Conference on Acoustics, Speech and Signal Processing, IEEE, 2010, pp. 349–352.
[35] P. Loiseau, P. Chevrel, M. Yagoubi, J.-M. Duffal, Broadband active noise control design through nonsmooth H_∞ synthesis, IFAC-Pap. 48 (14) (2015) 396–401.
[36] P. Strauch, B. Mulgrew, Active control of nonlinear noise processes in a linear duct, IEEE Trans. Signal Process. 46 (9) (1998) 2404–2412.
[37] L. Zhu, T. Yang, X. Li, L. Pang, M. Zhu, Active control of broadband noise inside a car using a causal optimal controller, Appl. Sci. 9 (8) (2019) 1531.
[38] Y. Zhou, Y. Yin, Q. Zhang, Active control of repetitive impulsive noise in a non-minimum phase system using an optimal iterative learning control algorithm, J. Sound Vib. 332 (18) (2013) 4089–4102.
[39] E. Sariyildiz, K. Ohnishi, A new solution for the robust control problem of non-minimum phase systems using disturbance observer, in: 2013 IEEE International Conference on Mechatronics, ICM, IEEE, 2013, pp. 46–51.
[40] L. Liu, F. Ripamonti, R. Corradi, Z. Rao, On the experimental vibroacoustic modal analysis of a plate-cavity system, Mech. Syst. Signal Process. 180 (2022) 109459.
[41] L. Cremer, M. Heckl, Structure-Borne Sound: Structural Vibrations and Sound Radiation at Audio Frequencies, Springer Science & Business Media, 2013.
[42] D. Chapelle, K.-J. Bathe, The Finite Element Analysis of Shells-Fundamentals, Springer Science & Business Media, 2010.
[43] K.S.A. Maamoun, S. Wrona, M. Pawelczyk, H.R. Karimi, Shaping of the frequency response of vibrating plates with openings for vibro-acoustic systems, Mech. Syst. Signal Process. 218 (2024) 111539.
[44] P. Lenk, G. Coult, Damping of glass structures and components, in: Challenging Glass Conference Proceedings, vol. 2, 2010, pp. 341–350.
[45] J.T. Du, W.L. Li, Z.G. Liu, H.A. Xu, Z.L. Ji, Acoustic analysis of a rectangular cavity with general impedance boundary conditions, J. Acoust. Soc. Am. 130 (2) (2011) 807–817.
[46] P. Morse, Sound wave in rooms, Rev. Modern Phys. 16 (1944) 100–117.
[47] R. Astley, G. Macaulay, J. Coyette, Mapped wave envelope elements for acoustical radiation and scattering, J. Sound Vib. 170 (1) (1994) 97–118.
[48] Y. Nakatsukasa, O. Sète, L.N. Trefethen, The AAA algorithm for rational approximation, SIAM J. Sci. Comput. 40 (3) (2018) A1494–A1522.
[49] A. Valera-Rivera, A.E. Engin, Aaa algorithm for rational transfer function approximation with stable poles, IEEE Lett. Electromagn. Compat. Pr. Appl. 3 (3) (2021) 92–95.
[50] I.V. Gosea, S. Gugercin, Data-driven modeling of linear dynamical systems with quadratic output in the AAA framework, J. Sci. Comput. 91 (1) (2022) 16.
[51] T.J. Moir, T.J. Moir, Rudiments of Signal Processing and Systems, Springer, 2022.
[52] G. Yu, Implementation of Kalman filter approach for active noise control by using MATLAB: Dynamic noise cancellation, 2024, arXiv preprint arXiv:2402.06896.
[53] G.L. Serra, Kalman Filters: Theory for Advanced Applications, BoD–Books on Demand, 2018.
[54] M.S. Aslam, M.A.Z. Raja, A new adaptive strategy to improve online secondary path modeling in active noise control systems using fractional signal processing approach, Signal Process. 107 (2015) 433–443.
[55] G.-y. Jin, T.-j. Yang, Y.-h. Xiao, Z.-g. Liu, A simultaneous equation method-based online secondary path modeling algorithm for active noise control, J. Sound Vib. 303 (3–5) (2007) 455–474.
[56] K.-A. Lee, W.-S. Gan, S.M. Kuo, Subband Adaptive Filtering: Theory and Implementation, John Wiley & Sons, 2009.

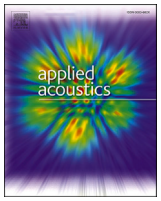
Applied Acoustics 239 (2025) 110823



Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust



Subjective perception analysis of active noise control algorithms in an encapsulated structure: An experimental study

Alkahf Aboutiman [a,b,*], Zulfi Rachman [b], Tin Oberman [b], Francesco Alletta [b], Jian Kang [b], Hamid Reza Karimi [a], Francesco Ripamonti [a]

[a] Politecnico di Milano, Department of Mechanical Engineering, Milan, Italy
[b] Institute for Environmental Design and Engineering, The Bartlett, University College London, London, UK

ARTICLE INFO

Keywords:
Active noise control
Encapsulated structure
Subjective listening
Deep learning algorithm

ABSTRACT

This study investigates the subjective perception of active noise control (ANC) performance, focusing on how individuals evaluate the noise reduction provided by different ANC algorithms. While the performance of the ANC algorithms has already been evaluated using objective metrics, this study aims to assess their effectiveness from a subjective perspective. In a simulated vehicle interior created using a noise box, two ANC algorithms were tested: the normalized least-mean-square (NLMS) algorithm and the hybrid selective fixed-filter active noise control normalized least-mean-square (SFANC-NLMS) algorithm. Participants were exposed to 27 stimuli, which combined three types of noise (motorcycle, street, and train), three sound pressure levels (55, 65, and 72 dB(A)), and three ANC conditions (no control, NLMS, and SFANC-NLMS). Subjective evaluations were collected using three indicators: perceived annoyance (PAY), perceived affective quality (PAQ), and perceived loudness (PLN). These metrics captured participants' impressions of the noise environment and the impact of noise control. The study is structured around three research questions (RQ1, RQ2, and RQ3), each addressing different aspects of ANC performance evaluation. In response to RQ1, the results demonstrated that the SFANC-NLMS algorithm outperformed NLMS in reducing perceived annoyance and loudness. Regarding RQ2, higher sound levels (72 dB) led to greater perceived annoyance, but sound level did not significantly alter the relationship between ANC algorithm type and perceived annoyance. Finally, in addressing RQ3, noise type influenced ANC effectiveness, with SFANC-NLMS showing more significant reductions in perceived annoyance compared to NLMS. Overall, the findings confirm that the SFANC-NLMS algorithm provides better noise reduction in encapsulated structures.

1. Introduction

Active noise control (ANC) is an effective technique for reducing unwanted sound or vibrations by applying the principle of sound wave superposition. The idea concept is generating a secondary signal, known as "anti-noise," which has the same amplitude as the unwanted noise but is phase-inverted, resulting in significant noise cancellation [1,2]. For this cancellation to occur, ANC systems must accurately predict the noise signal's amplitude and phase at a specific spatial point. While the amplitude may remain steady, the phase changes dynamically over time due to the physical behavior of sound waves in space [3,4]. ANC has found applications in numerous fields, such as personal audio devices, vehicles, aircraft, building acoustics, home appliances, industrial equipment, medical devices, and maritime environments. These technologies are particularly valuable in enclosed spaces like car interiors, airplane cabins, and train compartments [5,6], where passive noise reduction methods such as sound-absorbing materials—are traditionally employed [7]. However, at low frequencies, passive approaches require substantial material mass and volume to achieve significant noise reduction, which can be impractical [8,9]. In contrast, ANC systems are well-suited to addressing low-frequency noise challenges in such settings. Low-frequency noise in vehicles primarily originates from structural vibrations caused by the engine and tires [10,11]. Engine noise is typically concentrated below 500 Hz, while vibrations from tires often result in structure-borne noise below 1 kHz. Additionally, cabin resonances tend to amplify these low-frequency sounds, further contributing to the noise level. Road noise, generated by the interaction between vehicle tires and the road surface, represents another major source of noise in automo-

* Corresponding author.
E-mail address: alkahfahmat.aboutiman@polimi.it (A. Aboutiman).

https://doi.org/10.1016/j.apacoust.2025.110823
Received 2 March 2025; Received in revised form 11 April 2025; Accepted 9 May 2025
Available online 20 May 2025
0003-682X/© 2025 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

biles [12]. By addressing these challenges, ANC systems play a critical role in improving acoustic comfort and reducing noise pollution.

Traditional ANC systems typically rely on adaptive algorithms, such as the Filtered-X least-mean-square (FxLMS), to reduce noise due to their simplicity and robustness under standard conditions. However, these LMS-based algorithms exhibit slow convergence and limited tracking capability, making them less effective in handling dynamic or non-stationary noises commonly encountered in encapsulated structure where noise variability is high and user perception of noise reduction quality is critical [13,14]. Although other adaptive algorithms such as the Filtered-X Affine Projection (FxAP) and the Filtered-X Recursive Least Squares (FxRLS) offer improved convergence rates [15,16], their computational cost and sensitivity to modeling errors limit their applicability in real-time systems. Deep learning brings significant advancements to ANC by enabling models to adapt dynamically to complex, non-stationary noise patterns. Convolutional neural networks (CNNs) are effective in extracting robust features, eliminating the need for manually defined features that traditional algorithms rely on [17,18]. This automated feature extraction allows deep learning-based ANC systems to identify nuanced noise characteristics, offering better generalization across varying environments and noise types [19,20]. Building on these strengths, hybrid approaches like the selective fixed-filter active noise control (SFANC) combined with FxNLMS algorithm effectively combine deep learning with traditional adaptive control methods to overcome the limitations of each. In this setup, the CNN component dynamically selects the optimal pre-trained filter based on the noise type, while the FxNLMS algorithm adaptively fine-tunes the filter coefficients for real-time noise reduction. This synergy of SFANC-FxNLMS combines flexible filter selection with stable adaptive control, achieving faster response, enhanced noise reduction, and robustness in non-stationary environments such as vehicles and industrial settings [21].

Sound quality assessment has become a very active area of research, focusing on the subjective response that sounds (especially noise) evoke in humans, with the aim of achieving better sound design and noise enhancement techniques [22,23]. In particular, many studies have focused on the evaluation of sound quality in car interiors, examining different sound sources such as road noise, wind noise, and engine noise [24,25]. These studies highlight the complexity of the evaluation process, as each noise type has distinct characteristics, making it challenging to address all features simultaneously. Consequently, sound quality studies often focus on specific features, such as residual engine noise, which has been studied with active control. Results indicate that ANC is an effective tool for reducing low-frequency noise levels and improving acoustic comfort [26,27]. Studies have explored the perception of roughness in car engine noises using complex cepstrum analysis, underlining the importance of psychoacoustic parameters in designing sound profiles that align with user expectations [28]. While some experiments have been conducted in laboratory settings, further investigation is needed to fully understand ANC's impact on acoustic comfort inside vehicles. For instance, [29] designed an active tuning system to meet different engine sound quality requirements, and other methods have been explored for sound quality improvements with effectiveness demonstrated near virtual microphones [30,31] or error microphones [32,33]. Applied psychoacoustics is typically the foundation of sound quality evaluation [34,24], including in the automotive industry [35,36]. Psychoacoustic studies are often complemented by jury tests and mathematical models to estimate sound quality. Examples of sound quality studies on engine noise inside vehicles can be found in [37,38].

To address the research gap regarding the perception of active noise cancellation algorithms and their effects on noise perception in encapsulated structures, this study investigates the perceptual effects of active noise cancellation applied in a noise box simulating a vehicle interior. Specifically, this study compares a conventional NLMS algorithm with a modified version of SFANC-FxNLMS, called SFANC-NLMS, to investigate their influence on noise perception. The following research questions (RQ) are addressed:

- **RQ1.** To what extent do different ANC algorithms (NLMS and SFANC-NLMS) influence perceived annoyance across various noise types in a simulated vehicle environment?
- **RQ2.** Does sound level (55, 65, and 72 dB(A)) moderate the relationship between ANC algorithm type and perceived annoyance, altering the strength or direction of this association?
- **RQ3.** How do different noise sources (e.g., motorcycle, street, and train noises) interact with ANC algorithms and sound levels to shape perceived annoyance and overall acoustic comfort?

The paper is organized into five sections. Section 2 presents the context of this study by describing the environment of the control system and the algorithm used. Section 3 outlines the methodology employed to set up the experiment. Section 4 presents the results of the experiment. Section 5 provides a discussion of these results and answers the research questions. Finally, Section 6 concludes the paper.

2. Context of the study

2.1. Environment of control: noise box

The encapsulated structure, known as a noise box, is designed as a test platform for benchmarking vehicle interior noise. It is going to be a simplified interior noise investigation system coupled by a plate and a cavity. For example, this setup allows for the evaluation of materials, structures, and control strategies aimed at reducing noise inside a vehicle. The system must fulfill the following requirements:

- The system must be representative, providing access to panel, cavity, structure-borne noise, airborne noise, and noise control measures.
- The system should be straightforward to model and analyze, with a simple geometry and clearly defined boundary conditions.
- All geometric, material, and physical parameters must be specified.

A simplified plate-cavity system, modeled as a rigid box with one flexible panel, is selected to design a Noise-Box for testing vehicle interior noise control, based on key criteria for representativeness, ease of modeling, and defined parameters. Using a passenger car as a reference, the Noise-Box geometry reflects the relationship between the plate-cavity system and a vehicle cabin. Interior noise, which includes sounds from the engine, tires, intake, exhaust, wind, and other sources, enters through either structure-borne or airborne paths, as shown in Fig. 1. In this model, the cavity simulates the cabin, and the panel represents the vibrating car body: structure-borne noise arises when the panel is mechanically excited, and airborne noise occurs when external sounds pass through the panel, illustrating the coupled dynamics of the structure and enclosed acoustic field.

All control paths are calculated in this environment. The primary path refers to the path of the unwanted noise from the source to the error microphone. The secondary path is the path between the actuator, in this case, a loudspeaker, and the error microphone where the control signal is sent to cancel the primary noise. The reference path is the path from the noise source to the reference sensor, in this case, an accelerometer, which detects the primary noise characteristics and provides input to the control system for noise cancellation. The reference accelerometer is positioned at the top-left corner of the aluminum panel to avoid nodal points. The loudspeaker used as the actuator is fixed inside the cavity, facing the plate. The error microphone is located at the inside of the cavity. All of these control paths are computed through an experiment outlined in Section 3.2.1.

2.2. Traditional algorithm: NLMS algorithm

The NLMS algorithm is an adaptive active noise control (ANC) algorithm. Due to the non-minimum phase nature of our experimental environment, the FxNLMS algorithm, which typically relies on secondary

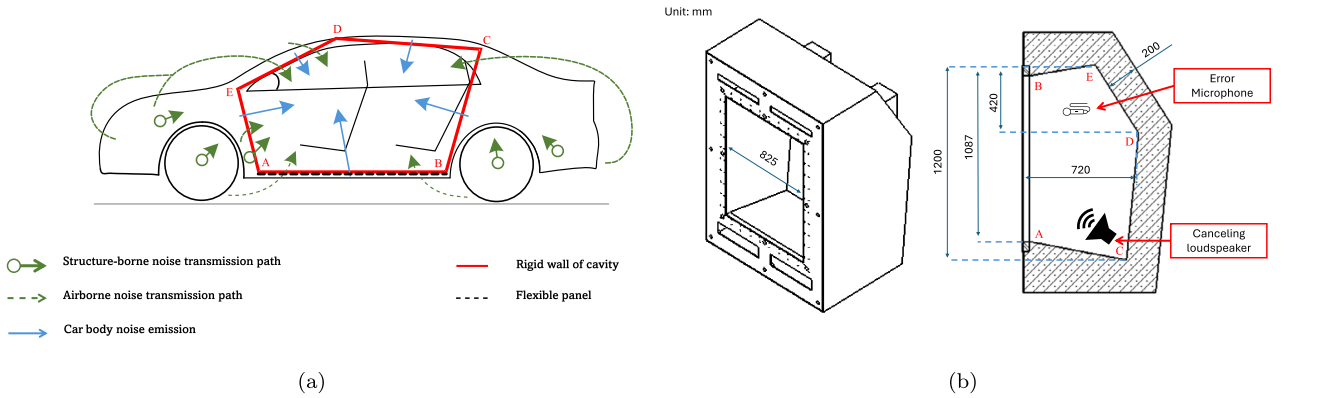


Fig. 1. The principle of noise in vehicle interiors and the panel-cavity system (a) and schematic diagram of the Noise Box for noise mitigation measures (b) [39].

path estimation, was found to be unsuitable for this study. A system is said to be non-minimum phase when it presents unstable zeros, i.e., zeros located outside the unit circle in the z-domain. These zeros result in an initial system response that moves in the opposite direction of the steady-state behavior, making control more challenging and often introducing phase-related delays. Non-minimum phase systems often lead to instability or reduced performance in ANC applications [40,41]. In practice, secondary path estimation is subject to modeling errors, which degrade control performance, particularly in experimental environments. Based on these considerations and the results from the NLMS algorithm, the FxNLMS algorithm was forgone in favor of the simpler NLMS algorithm, avoiding the complexities of secondary path estimation.

The control system is based on a feedforward architecture, with the input signal vector $x(n) = [x(n), x(n-1), \ldots, x(n-N+1)]$ and the signal after the reference path $r(n)$ drive the NLMS control output $y(n)$:

$$y(n) = w^T(n)r(n) \tag{1}$$

where $w(n)$ represents the adaptive filter coefficient vector. The residual error signal, $e(n)$, defined as:

$$e(n) = d(n) - s(n) * (w^T(n)r(n)) \tag{2}$$

This error is minimized using the gradient descent method with step size μ. The NLMS update rule is:

$$w(n+1) = w(n) + \mu e(n)\frac{r(n)}{r^T(n)r(n) + \epsilon} \tag{3}$$

where T denotes the transpose operation, and ϵ is a small constant (set to 2.2204×10^{-16}) to avoid numerical instability during weight updates [42].

2.3. Novative algorithm: SFANC-NLMS algorithm

The hybrid SFANC-NLMS algorithm combines adaptive filtering with a dynamic filter selection mechanism that adjusts the control filter based on the type of noise. The algorithm operates with a set of pre-trained control filters, each optimized for different noise profiles. A 1D convolutional neural network (CNN) processes the incoming noise and selects the most appropriate control filter by outputting the index of the filter best suited to the current noise conditions. Once the optimal filter is chosen, the NLMS algorithm continuously updates its coefficients to minimize the error signal further. This adaptive adjustment occurs at the sampling rate, allowing the control filter to be fine-tuned in real time. By integrating the filter selection capabilities of the CNN with the adaptability of the NLMS algorithm, the system efficiently adapts to varying noise conditions. This approach is presented in more detail in [43].

The simulation parameters are as follows: The ANC system operates in a single-channel configuration with a sampling rate of 16 kHz, and the control filter length is set to 4096 taps. The system utilizes state-space coefficients for the reference, primary, and secondary paths. Seven pre-trained control filters were used, corresponding to seven different white noise signals with the following frequency ranges: 20–2000 Hz, 20–1000 Hz, 1000–2000 Hz, 20–500 Hz, 500–1000 Hz, 1000–1500 Hz, and 1500–2000 Hz. The NLMS algorithm was employed to derive the optimal control filters for these noises, with a stepsize of 0.0001.

3. Methodology

3.1. Study site, participants and administration

The experiment was conducted in a study room at Central House, chosen for its quiet environment, where participants used paper and pencil to complete the tasks as presented in Fig. 2. Formal ethical approval was approved by BSEER Local Research Ethics Committee at University College London prior to participant recruitment and the commencement of the experiment. Each participant was informed about the study context through a Participant Information Sheet before taking part, and consent was obtained via a form at the beginning of the experiment.

A total of 35 participants, including 16 men and 19 women, aged 23 to 62 (mean age around 31), were recruited. Two participants were deemed unreliable due to a Spearman coefficient of 0.8 and a p-value over 20%.

3.2. Stimuli

3.2.1. Experimental setup for stimuli design in the noise box

The stimuli design for the jury test relies on the experimental setup of the noise box, as introduced in Section 2.1. This setup includes a sealed cavity, where the opening is closed with an aluminum plate to isolate and focus acoustic interactions. The noise box is equipped with components to extract the frequency response functions (FRFs) of the reference, primary, and secondary paths, to characterize the acoustic behavior of the system. Fig. 3 illustrates this configuration, showcasing the internal and external arrangement.

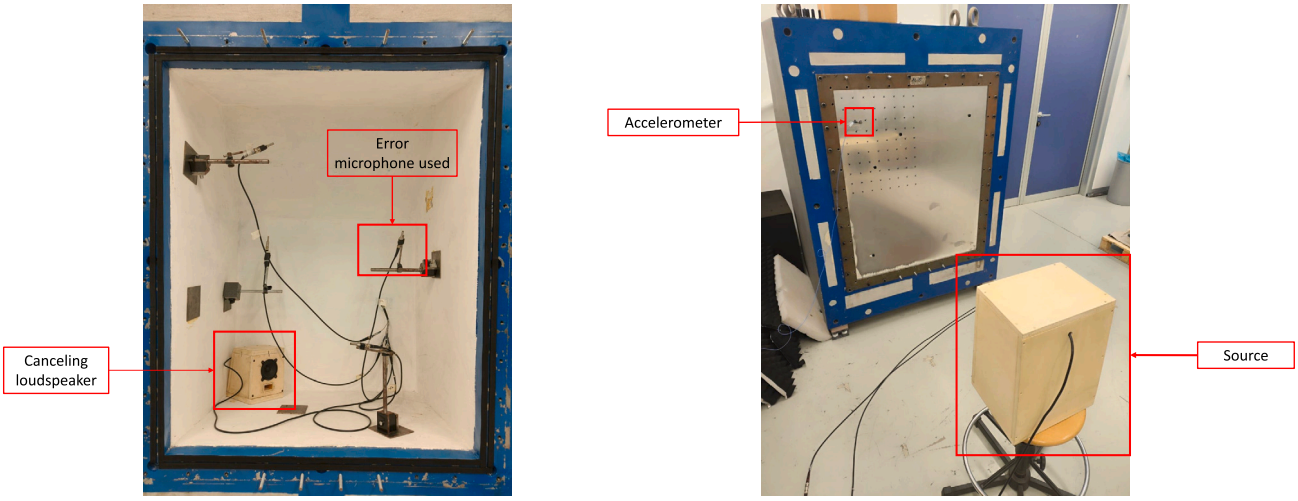
Two speakers are utilized in this setup. An external speaker simulates disturbances outside the cavity, while an internal speaker generates the “anti-noise” signal, it is called a canceling loudspeaker. Both speakers are driven by random noise signals generated by a computer and transmitted through the data acquisition system. Inside the cavity, four microphones are placed to capture acoustic responses, though only one microphone is employed for this single-channel control algorithm experiment. An accelerometer, positioned at the top left corner of the panel to avoid nodal points, acts as the reference sensor by capturing noise vibrations before they enter the cavity.

The measurement equipment includes a PCB 130E20 microphone with a sensitivity of 37.37 mV/Pa, calibrated using a Bruel & Kjaer 4231



(a) (b)

Fig. 2. Illustration of Participant Participation. (a) shows how the participant filled out the questionnaire, and (b) depicts the room where the experiment took place.



(a) (b)

Fig. 3. Images of the experimental set-up from inside (a) and outside (b).

calibrator providing a reference frequency of 1 kHz and a sound pressure level of 94 dB. An accelerometer, the PCB 333B30 model with a sensitivity of 99.6 mV/g (10.16 mV/m/s^2), is connected to a PCB Piezotronics model 483C series signal conditioner, interfacing with a National Instruments NI cDAQ-9178 data acquisition system. These devices ensure accurate measurements of both acoustic and vibratory signals.

The FRFs were derived in two phases. First, the external speaker was activated to measure the reference path, with the input being the voltage signal to the speaker and the output being the accelerometer signal (m/s^2), resulting in an FRF magnitude in m $\cdot$ V/s^2. Next, the internal speaker was used to compute both the primary and secondary path FRFs. The primary path was calculated using the external speaker's input voltage and the cavity microphone signal as output (Pa/V), while the secondary path was derived similarly but with the internal speaker as the input source. All FRFs were computed over a frequency range of 100 Hz to 2000 Hz, with a sampling rate of 16,000 Hz. Additionally, the acoustic characteristics of the laboratory, influenced by surrounding materials, shaped the behavior of the sound field. These factors were accounted for when analyzing the experimental data.

To replicate real-world scenarios, three types of noise, motorbike, train, and street sounds, were selected from a publicly available dataset [44]. These noises were filtered through the state-space model derived from the FRFs of the noise box, ensuring realistic acoustic behavior. The FRFs also enabled the derivation of the state-space coefficients needed to model the acoustic dynamics of the noise box, a critical step for implementing active noise control. In the experimental design, three ANC conditions (off, NLMS, and SFANC-NLMS) were tested with the three noise types at three sound pressure levels (55 dB(A), 65 dB(A), and 72 dB(A)), resulting in $3 \times 3 \times 3 = 27$ stimulus combinations. Additionally, a train noise condition at 65 dB(A) was included as a control to as-

Table 1
Mean and standard deviation of the measured LAeq levels for each noise type and control condition. The first line is the level on the left side of the HATS and the second one is the right sides.

Noise Type	Control: OFF			Control: NLMS			Control: SFANC-NLMS		
	72 dB(A)	65 dB(A)	55 dB(A)	72 dB(A)	65 dB(A)	55 dB(A)	72 dB(A)	65 dB(A)	55 dB(A)
Train	73.4 ± 0.3 73.0 ± 0.3	67.1 ± 0.5 67.0 ± 0.4	56.4 ± 0.6 55.4 ± 0.6	73.0 ± 0.3 73.4 ± 0.6	67.3 ± 0.4 67.9 ± 0.6	56.2 ± 0.1 56.4 ± 0.1	68.3 ± 0.4 69.2 ± 0.5	63.1 ± 0.3 63.1 ± 0.4	50.2 ± 0.2 50.0 ± 0.2
Moto	70.3 ± 0.3 70.2 ± 0.2	66.3 ± 0.2 66.3 ± 0.3	56.4 ± 0.3 56.5 ± 0.3	69.3 ± 0.1 68.9 ± 0.3	63.9 ± 0.1 63.4 ± 0.3	54.8 ± 0.3 54.7 ± 0.3	62.8 ± 0.3 62.7 ± 0.3	58.1 ± 0.3 58.5 ± 0.4	48.8 ± 0.4 48.5 ± 0.3
Street	72.0 ± 0.7 71.4 ± 0.6	64.3 ± 0.4 64.3 ± 0.3	57.3 ± 0.4 57.5 ± 0.4	72.2 ± 0.3 72.1 ± 0.3	64.1 ± 0.4 64.2 ± 0.3	56.9 ± 0.3 57.4 ± 0.3	67.3 ± 0.4 67.4 ± 0.4	61.8 ± 0.3 61.8 ± 0.4	54.6 ± 0.3 54.6 ± 0.3



(a) (b)

Fig. 4. Interior view of the Audio Lab with the HATS setup (a), and the HATS setup during the calibration (b).

sess test–retest reliability, bringing the total to 28 unique stimuli for the study.

3.2.2. Calibration

The calibration was conducted in an Audio Lab, using a Head and Torso Simulator (HATS) HEAD acoustics HMS II.3 LN HEC with Sennheiser HD 650 headphones. Before calibrating the stimuli, the HATS was initially calibrated with a Larson Davis CAL250 sound level calibrator, emitting a 251.2 Hz tone at 114 dB. Each stimulus, played through a computer, was calibrated individually. Volume settings were adjusted to achieve the desired noise levels of 55 dB(A), 65 dB(A), and 72 dB(A). These levels were calibrated without the application of active noise control. For stimuli with control applied, the volume parameters were adjusted according to the corresponding noise level. ArtemiS SUITE 12 software was used for the calibration, along with the labO2-V1 playback equalizer and the HDA IV headphone amplifier. To account for variability in noise levels due to headphone positioning, each stimulus was calibrated three times with the headphones repositioned on the HATS between each measurement to capture uncertainty. The results of the measured noise levels after calibration are shown in Table 1, and photographs of the experiment setup are provided in Fig. 4.

3.3. Experiment design

The participants were first required to provide basic information: gender and age, along with a pre-test assessment of (1) individual noise sensitivity (INS) via the Weinstein Sensitivity to Sound Scale, 5-item version (WSSN-5) [45]; (2) a baseline noise annoyance (BNA) based on ISO/TS 15666 [46] using the version of the questionnaire presented in [47]; and (3) the Perceived Stress Scale (PSS-4) [48]. The pre-test questionnaires were selected to assess potential confounding factors such as stress and mood on soundscape perception [49,50]. The pre-test assessment questionnaire is presented in Table A.1 in Appendix A.

After completing the basic data and pre-test assessment, a brief training session was conducted to familiarize participants with the stimuli and questionnaire. During training, the participants were first exposed to the loudest stimulus, followed by the quietest stimulus, and then a training noise (a train noise at 65 dB(A)). After hearing the training noise, participants were required to answer questions. Stimuli were played through a sound sequence design using the software FL Studio.

For test–retest reliability, an unfiltered 65 dB(A) train noise was used as the first and last stimulus for each participant, without replicating the noise box behavior. The other 27 stimulus combinations were presented in random order to all participants. Participants were allowed to listen to each stimulus as much as they wanted, with an experimenter available to replay stimuli if needed. All stimuli presented to the participants were recorded and played back.

Each stimulus was evaluated based on (1) perceived annoyance (PAY), (2) perceived affective quality (PAQ), and (3) perceived loudness (PLN). In contrast to the pre-test assessment of annoyance in the absence of acoustic stimuli, the assessment of PAY referred specifically to the noise in the presented stimulus. The question asked: *Thinking about the noise you just heard, how much does the noise bother, disturb, or annoy you?* Responses were divided into 5 items, ranging from “Not at all” to “Extremely” [46].

The PAQ (Perceived Affective Quality) attributes proposed in ISO/TS 12913-2 were adopted to structure the evaluation of affective responses to the indoor soundscape as modified by the stimuli. The PAQ consists of 8 attributes (i.e., eventful, vibrant, pleasant, calm, uneventful, monotonous, annoying, chaotic), forming an octant circumplex model (International Organization for Standardization, 2019 [51]). Participants were asked to visually position a mark on a 2D map representing these attributes. The question asked: *Thinking about the sound you just heard, where would you place the surrounding sound environment on this scale?* This method allows participants to express their affective response to the sound environment, accounting for both its emotional tone and complexity.

To assess loudness (PLN), a relative magnitude estimation method was adopted. Participants were instructed to estimate the loudness of the stimulus under test (SUT). A numerical score was assigned to each stimulus based on its perceived loudness. The question asked: *How loud would you say the sound environment is?* The scale ranged from 0 to 100, allowing for a nuanced assessment of loudness perceived by the participant [52]. The noise evaluation questionnaire is presented in Table A.2 in Appendix A.

3.4. Data analysis

The data were processed to obtain the mean scores and standard deviations for each indicator: INS, PSS-4, and BNA. For the INS and PSS-4 indicators, each participant's score was calculated by summing the scores for all responses given to the questions and dividing this sum by the total number of questions in the respective questionnaire. Specifically, the INS questionnaire consists of 5 questions, and the PSS-4 questionnaire consists of 4 questions. The global mean score for each indicator was then calculated by averaging the individual participant scores. Standard deviations were computed to assess the variability of participants' scores around the mean.

For the BNA indicator, the data were organized by noise type (e.g., traffic, airplane, train, etc.). Each participant's score for each noise type was calculated in the same manner, by summing the responses to the relevant questions and dividing by the number of questions for that noise type. The mean and standard deviation were then calculated for each noise type based on the participants' scores. These means and standard deviations provide an overview of participants' perceptions regarding different noise types and indicators.

A three-way repeated measures ANOVA (3WR-ANOVA), followed by a post-hoc Tukey HSD test, was conducted to evaluate perceived annoyance (PAY), with the numerical scale of annoyance as the dependent variable and three independent variables (i.e., noise type, noise levels, and ANC condition). Prior to conducting the analysis, we verified that the assumptions of normality, homogeneity of variances, and sphericity were met. The normality of residuals was assessed using the Shapiro-Wilk test [53], homogeneity of variances was tested with Levene's test [54], and sphericity was tested using Mauchly's test [55]. All assumptions were satisfied, allowing us to proceed with the 3WR-ANOVA. The circumplexity of the PAQ attributes was examined to assess the generalizability of the PAQ model. Furthermore, a three-way repeated measures permutational multivariate analysis of variance (3WR-PERMANOVA) was conducted using a distance matrix as input. This method was chosen because the PAQ data did not follow a normal distribution, as indicated by Mardia's multivariate normality tests [56]. The analysis considered ISOPL, the x-coordinate of the participant's mark on the 2D PAQ map, and ISOEV, the y-coordinate, as the dependent variables. The independent variables included noise levels, control conditions, and noise types. Post-hoc pairwise comparisons were performed using the pairwise.adonis() function, based on the ADONIS (permutational multivariate analysis of variance) method, to examine the differences between factor levels across all combinations of noise types, control conditions, and noise levels. This procedure was employed to identify significant pairwise differences between the groups following the main analysis.

For perceived loudness (PLN), a three-way repeated measures ANOVA (3WR-ANOVA), followed by a post-hoc Tukey HSD test, was conducted to evaluate perceived loudness, with the numerical scale of loudness as the dependent variable and three independent variables (i.e., noise type, noise levels, and ANC condition). Similar to the PAY analysis, various tests were performed on the data to check if all assumptions were satisfied in order to use the 3WR-ANOVA.

4. Results

4.1. Pre-test assessment

The mean value for the INS indicator is 6.49 with a standard deviation of 2.20, which indicates that the participants on average experience a medium to high level of noise pollution in their environment. As the INS scale ranges from 1 to 10, this value is in the upper mid-range of the scale and indicates a medium to high sensitivity to noise among the participants. In contrast, the mean value for PSS-4, which measures perceived stress, is 3.12 with a standard deviation of 0.75. As the scale for this indicator ranges from 0 to 4, this value indicates that the participants experience a medium level of stress on average, with some variability in the responses, but without reaching an extreme level of stress.

The mean scores for each type of noise, as measured by the BNA, vary slightly. The highest mean score was recorded for road traffic noise with a value of 2.37, suggesting that, on average, this noise type is perceived as somewhat disturbing. The lowest mean score was recorded for animals noise with a value of 1.51, indicating that participants generally find animal-related noises less disturbing. The other noise types, including airplane, train, children, and other people have mean scores ranging from 1.86 to 2.09, reflecting a moderate level of disturbance for these sounds. In terms of standard deviations, road traffic and airplane have the highest variability, with standard deviations of 1.14 and 1.19, respectively, indicating that responses to these noises are more diverse. The animals and other people categories show the lowest variability, with standard deviations of 0.95, suggesting more consistent perceptions of these noise types across participants. Overall, these results show moderate variability in how different noise types are perceived by participants.

4.2. Perceived annoyance

The results of the three-way repeated ANOVA (3WR-ANOVA) indicate that the control condition, noise type, and noise level all have significant effects on the perceived annoyance of participants, with some interactions between these factors revealing more complex relationships. In this analysis, the F-value reflects the ratio of variance explained by each factor to the unexplained variance (residual error). A higher F-value indicates a greater influence of the corresponding factor on perceived annoyance. The p-value represents the probability that the observed effect occurred by chance, with values below the significance threshold (typically 0.05) indicating a statistically significant effect. In this study, all p-values have been adjusted using the False Discovery Rate (FDR) correction to account for multiple comparisons. The control condition shows a particularly strong effect ($F = 44.48, p < .0001$), suggesting that different noise control strategies significantly influence how participants perceive noise. Similarly, noise level has a major impact ($F = 161.4, p < .0001$), with higher levels of noise leading to greater perceived annoyance. Noise type also contributes significantly ($F = 5.336, p = .007$), indicating that different types of sounds (e.g., street, train, or motorbike noise) have varying impacts on participants' experiences. When examining interactions, the combination of control condition and noise type is significant ($F = 4.859, p = .0015$), suggesting that the effectiveness of noise control methods differs depending on

Table 2
Results of the three-way repeated measures ANOVA (3WR-ANOVA) for PAY.

Comparison	F	p-value (FDR)
Noise Types	5.33	0.007 *
ANC	44.48	< 0.0001 ****
Noise Levels	161.4	< 0.0001 ****
Noise Types × Control Conditions	4.86	0.0015 **
Noise Types × Noise Levels	7.93	< 0.00001 ****
Control Conditions × Noise Levels	1.83	0.126
Noise Types × Control Conditions × Noise Levels	3.47	0.0008 ***

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

the type of noise present. This highlights that noise control strategies may need to be tailored to specific noise sources to achieve optimal results. Additionally, the interaction between noise type and noise level is significant ($F = 7.926, p < .0001$), implying that the combined characteristics of the noise (both type and intensity) play an important role in the perceived annoyance of the participant. However, the interaction between control condition and noise level ($F = 1.831, p = .126$) does not show significant effects, meaning that noise control effectiveness is relatively consistent across different noise levels. Finally, the three-way interaction between control condition, noise type, and noise level is significant ($F = 3.47, p = .0008$), indicating that these factors do not operate independently but interact in ways that can influence noise perception in more complex ways. These findings underscore the importance of considering multiple factors together when designing noise control strategies, as their combined effects can vary depending on the specific noise characteristics. Results of the 3WR-ANOVA are findable in Table 2.

Following results obtained with the post hoc Tukey HSD, significant differences in the perceived noise annoyance were observed across different noise types, control conditions, and sound levels. The effect size d complements the p-value by quantifying the magnitude of the difference between two conditions. This allows a better understanding of the practical significance of observed differences beyond mere statistical significance. Regarding noise types, the comparison between motorbike and street shows a highly significant difference with a moderate effect size noted d ($p = .005, d = -0.27$), indicating that motorbike noise is perceived differently from street noise. However, the comparisons between motorbike and train ($p = .370, d = -0.11$) and street and train ($p = .147, d = .15$) show no significant differences, suggesting relative similarity between these noise types in terms of perceived annoyance. For the control conditions, all comparisons are extremely significant ($p < .0001$ or $p < .001$), with effects ranging from moderate to strong. SFANC-NLMS provides a notable improvement compared to No Control ($d = -0.53$) and NLMS ($d = -0.31$), confirming the superior effectiveness of this new algorithm. The No Control condition differs significantly from NLMS ($d = 0.22$), although the effect size is smaller, highlighting some improvement with NLMS. The interactions between noise types and control conditions reveal marked differences. For instance motorbike with SFANC-NLMS significantly differs from street with No Control ($p < .0001, d = -0.98$), while moderate to strong effects are observed for other combinations. These results underscore the importance of interactions between noise type and algorithm effectiveness. Finally, the three-way interactions between noise levels, noise types, and control conditions reveal interesting patterns. For example, for motorbike noise at 55 dB, SFANC-NLMS is perceived as substantially more effective than No Control at 72 dB ($d = 2.20$), indicating that the improvement provided by the algorithm can offset a higher noise level. These interactions demonstrate that the perceived effectiveness of the algorithms varies depending on the specific combination of noise type, sound level, and control condition, highlighting the need to tailor noise reduction strategies to these factors. Results of the post hoc Tukey HSD are findable in Table 3.

Table 3
Post hoc Tukey HSD results for significant comparisons in the PAY study.

Comparison	p-value	Effect
Moto vs. Street	0.005	-0.27
Moto vs. Train	0.37	-0.1111
Street vs. Train	0.15	0.16
SFANC-NLMS vs. No Control	< 0.0001 ****	-0.53
SFANC-NLMS vs. NLMS	< 0.0001 ****	-0.31
No Control vs. NLMS	0.0006 ***	0.22
55 dB vs. 65 dB	< 0.0001 ****	-0.75
55 dB vs. 72 dB	< 0.0001 ****	-1.43
65 dB vs. 72 dB	< 0.0001 ****	-0.69
Moto (SFANC-NLMS) vs. Street (No Control)	< 0.0001 ****	-0.98
Street (No Control) vs. Train (SFANC-NLMS)	< 0.0001 ****	+0.77
Moto (SFANC-NLMS) vs. Moto (No Control)	< 0.0001 ****	-0.60
Moto (55 dB) vs. Moto (72 dB)	< 0.0001 ****	-1.54
Moto (55 dB) vs. Street (72 dB)	< 0.0001 ****	-1.75
Moto (55 dB) vs. Train (72 dB)	< 0.0001 ****	-1.71
Train (55 dB) vs. Street (72 dB)	< 0.0001 ****	-1.65
Train (55 dB) vs. Train (72 dB)	< 0.0001 ****	-1.61
SFANC-NLMS (55 dB) vs. SFANC-NLMS (72 dB)	< 0.0001 ****	-1.51
SFANC-NLMS (55 dB) vs. No Control (72 dB)	< 0.0001 ****	-1.88
SFANC-NLMS (55 dB) vs. NLMS (72 dB)	< 0.0001 ****	-1.79
NLMS (55 dB) vs. No Control (72 dB)	< 0.0001 ****	-1.62
NLMS (55 dB) vs. NLMS (72 dB)	< 0.0001 ****	-1.53
SFANC-NLMS (Moto, 55 dB) vs. No Control (Street, 72 dB)	< 0.0001 ****	2.20
SFANC-NLMS (Moto, 55 dB) vs. NLMS (Train, 72 dB)	< 0.0001 ****	2.03
NLMS (Moto, 55 dB) vs. No Control (Street, 72 dB)	< 0.0001 ****	2.11
SFANC-NLMS (Train, 55 dB) vs. No Control (Street, 72 dB)	< 0.0001 ****	2.23

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

4.3. Perceived affective quality

4.3.1. Circumplex circles

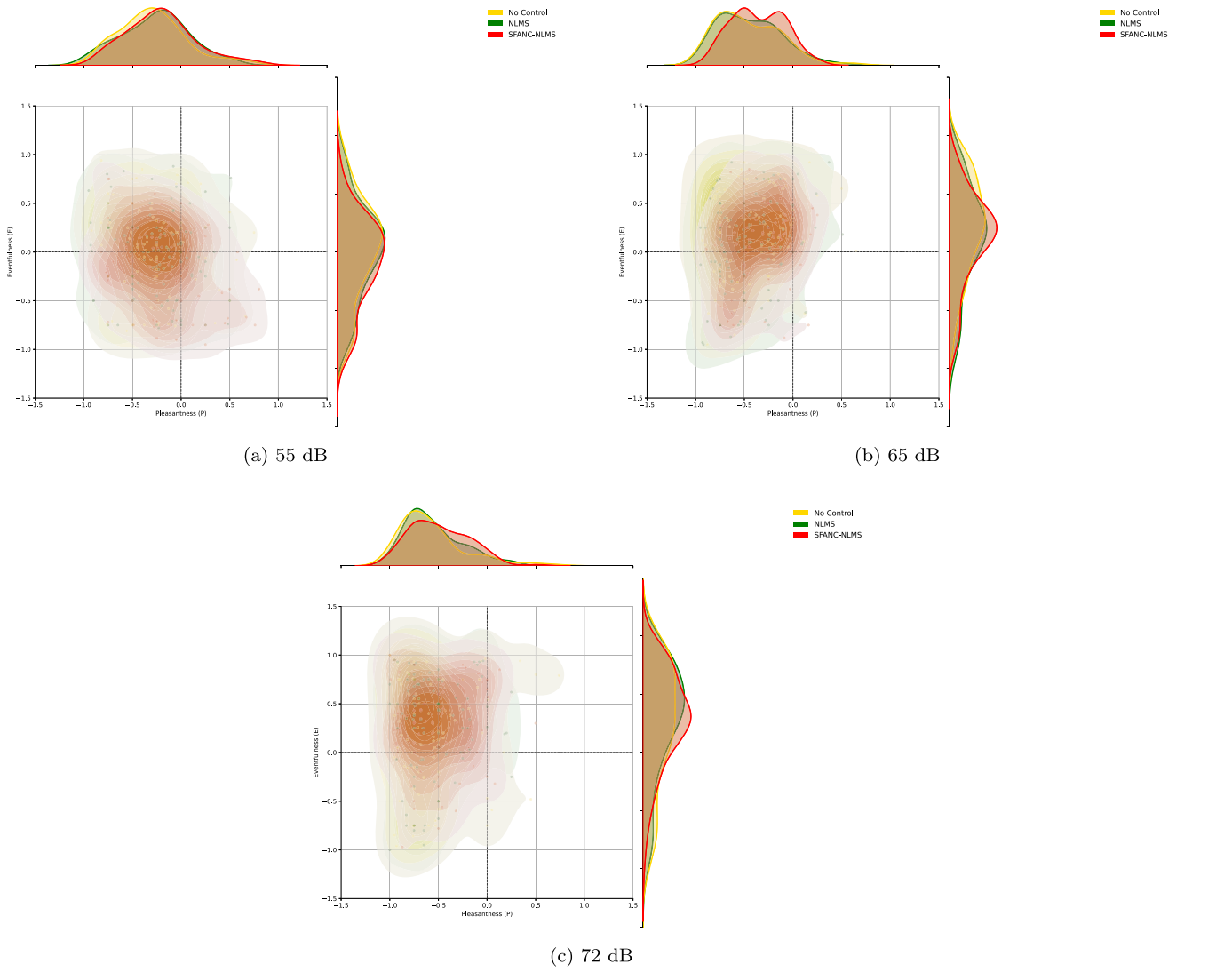
Regarding the distribution of points on the circumplex circles, we decided to merge results obtained with different types of noise to highlight the influence of the control condition. Fig. 5 displays plots of circumplex circles at various noise levels (55, 65, 72 dB(A)). Noises are generally perceived as chaotic, especially at higher sound levels. This perception can be attributed to two main factors. First, filtering each noise to replicate the behavior of the noise box amplifies the aluminum plate's acoustic response, adding a distinctive resonant character to the stimuli. Second, the original noises themselves are inherently eventful, characterized by dynamic and varied sound patterns. These two aspects contribute to an increased perception of eventfulness and a reduced sense of pleasantness, particularly at higher intensities.

As sound levels increase, the distribution of responses shifts toward the upper-left quadrant of the circumplex space—indicative of a more chaotic and unpleasant perceptual evaluation. Naturally, stimuli with lower sound levels are generally perceived as more pleasant, aligning with the typical relationship between sound intensity and subjective annoyance.

Regarding the control condition, the shape and position of the density clouds visible through the kernel density estimates on the x (pleasantness) and y (eventfulness) axes indicate that applying SFANC-NLMS makes the noise perceived as more pleasant and less eventful across all noise levels. This pattern suggests that the algorithm improves the affective quality of the soundscape, shifting it toward the bottom-right quadrant, which corresponds to calmer and more pleasant environments.

4.3.2. Three way repeated PERMANOVA

The results of the three-way repeated-measures PERMANOVA revealed several significant effects concerning the influence of the factors noise type, control condition, and noise level on the dependent variables ISOPL and ISOEV, as well as their interactions. The effect of noise type was highly significant ($F = 11.17, p = .001$), indicating that the



(a) 55 dB (b) 65 dB

(c) 72 dB

Fig. 5. Density plots of eventfulness (ISOEV) as a function of pleasantness (ISOPL). Plots are shown separately for (a) 55 dB(A) (b) 65 dB(A) and (c) 72 dB(A).

type of noise strongly influences the dependent variables, independent of the other factors. This confirms that different noise types lead to distinct perceptual differences. Similarly, the effect of control condition was significant ($F = 9.35, p = .001$), suggesting that different control strategies influence participants' perceived affective quality. While this effect is notable, it is slightly lower than that of noise type. The noise level effect was also highly significant ($F = 13.50, p = .001$), demonstrating that increasing noise intensity (55, 65, 72 dB) significantly impacts participants' perception. This confirms that noise level is a key determinant of perceived affective quality. Unlike the initial analysis, the interactions between the factors also reached significance. The interaction between noise type and control condition ($F = 10.01, p = .001$) suggests that the effect of noise type varies depending on the applied control strategy. The interaction between noise type and noise level ($F = 14.00, p = .001$) reveals that the impact of noise type is modulated by its intensity. Additionally, the interaction between control condition and noise level ($F = 12.07, p = .001$) suggests that the effectiveness of control strategies depends on noise intensity. Finally, the three-way interaction between noise type, control condition, and noise level was also significant ($F = 10.43, p = .001$). This indicates that the combined effect of all three factors significantly influences participants' perception, meaning that noise type, control conditions, and noise level interact in a complex way to shape perceived affective quality. Results of the three-way repeated PERMANOVA are presented in Table 4.

Table 4
Results of the three-way repeated measures PERMANOVA (3WR-PERMANOVA) for PAQ.

Comparison	F	p-value
Noise Types	11.17	0.001 **
ANC	9.35	0.001 **
Noise Levels	13.50	0.001 **
Noise Types × Control Conditions	10.01	0.001 **
Noise Types × Noise Levels	14.00	0.001 **
Control Conditions × Noise Levels	12.07	0.001 **
Noise Types × Control Conditions × Noise Levels	10.43	0.001 **

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

The results of the post hoc pairwise revealed significant perceptual differences across several conditions. Specifically, the comparison between the motorbike and street noise types showed a substantial difference in affective quality ($p = .003, d = -0.120$), indicating that participants found the motorbike noise to be perceived as more unpleasant or intrusive compared to the street noise. Similarly, the comparison between street and train also revealed a significant difference ($p = .003, d = .090$), suggesting that the street noise was perceived more positively compared to the train noise, with an increased sense of unpleasantness for the train noise. In contrast, the comparison between motorbike and train noise types ($p = .021, d = -0.030$) did not reveal a

Table 5
Post-hoc pairwise comparisons results using the ADONIS method for significant comparisons in the PAQ study.

Comparison	p-value	Effect
Moto vs. Street	0.003 **	-0.12
Moto vs. Train	0.021 *	-0.03
Street vs. Train	0.003 **	0.09
NLMS vs. No Control	1.000	-0.03
NLMS vs. SFANC-NLMS	0.056	0.168
No Control vs. SFANC-NLMS	0.011 *	0.03
55 dB vs. 65 dB	0.003 **	-0.09
55 dB vs. 72 dB	0.003 **	-0.24
65 dB vs. 72 dB	0.003 **	-0.153
Moto (NLMS) vs. Street (No Control)	0.036 *	-0.2
Street (No Control) vs. Train (No Control)	0.036 *	0.19
Street (No Control) vs. Moto (SFANC-NLMS)	0.036 *	0.23
Street (No Control) vs. Train (SFANC-NLMS)	0.04 *	0.172
Moto (55 dB) vs. Street (72 dB)	0.036 *	-0.36
Street (72 dB) vs. Train (55 dB)	0.036 *	0.37
NLMS (55 dB) vs. No Control (72 dB)	0.036 *	-0.26
NLMS (55 dB) vs. NLMS (72 dB)	0.036 *	-0.25
No Control (72 dB) vs. SFANC-NLMS (55 dB)	0.036 *	0.26
SFANC-NLMS (55 dB) vs. NLMS (72 dB)	0.036 *	-0.26

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

significant perceptual difference after correction, suggesting that these two noise types were perceived similarly in terms of affective quality by the participants. Regarding control conditions, a significant difference was observed between the No Control and SFANC-NLMS conditions ($p = .011, d = .033$). This indicates that the SFANC-NLMS algorithm provided some improvement in the perceived affective quality of the noise environment, although the effect size was relatively small. In terms of noise levels, several significant differences were found. The comparison between 55 dB and 65 dB ($p = 0.003, d = -0.090$), 55 dB and 72 dB ($p = .003, d = -0.240$), and 65 dB and 72 dB ($p = .003, d = -0.153$) all revealed a significant impact on the perceived affective quality. The higher the noise level, the more negative the affective perception became, with 72 dB generally being perceived as the most unpleasant or disturbing, followed by 65 dB and 55 dB. Several other comparisons, such as those involving noise types (Moto vs. Street, Street vs. Train) and noise levels (55 dB vs. 72 dB, 65 dB vs. 72 dB), also showed statistically significant differences. These results indicate that participants were able to detect clear and meaningful differences in how different types of noise and noise levels affected their emotional response, providing valuable insights into the subjective affective perception of different acoustic environments. The detailed results of the post hoc comparisons are presented in Table 5.

4.4. Perceived loudness

The results of the three-way repeated measures ANOVA revealed that the control condition, noise type, and noise level all have significant effects on the perceived loudness of participants. A highly significant effect was found for noise types ($F = 9.739, p = .00019$), suggesting that the different categories of noise (e.g., motorcycle, street, train) have a substantial impact on the perceived loudness. This indicates that perceived loudness varies depending on the type of noise, with some types being perceived as more intense or disturbing than others. The main effect of noise control conditions was extremely significant ($F = 41.29, p < .0001$), highlighting the strong influence of noise control algorithms on the perceived loudness. The results indicate that these algorithms significantly reduce the perceived loudness of the noise, regardless of its type, showing their effectiveness in mitigating noise. Similarly, noise level has a major impact ($F = 63.1, p < .0001$), confirming that higher noise levels (55 dB, 65 dB, 75 dB, etc.) lead to stronger

Table 6
Three-way repeated ANOVA results showing F-values, p-values, and significance levels for main effects and interactions for the PLN.

Comparison	F	p-value
Noise Types	9.74	< 0.001 ***
ANC	41.29	< 0.0001 ****
Noise Levels	63.1	< 0.0001 ****
Noise Types × Control Conditions	5.171	0.0007 ***
Noise Types × Noise Levels	7.93	< 0.0001 ****
Control Conditions × Noise Levels	1.11	0.36
Control Conditions × Noise Types × Noise Levels	2.65	0.008 **

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

perceptions of loudness, with louder sounds being perceived as more disturbing. In terms of interactions, a significant interaction between noise types and control conditions was observed ($F = 5.171, p = .000658$), indicating that the effect of noise types on perceived loudness depends on the control algorithm used. A significant interaction between noise types and noise levels was also found ($F = 7.934, p < .0001$), showing that the effect of noise levels on perceived loudness differs according to the type of noise. On the other hand, the interaction between control conditions and noise levels was not significant ($F = 1.107, p = .356$), suggesting that noise levels did not significantly affect the effectiveness of the control algorithms in terms of perceived loudness. Lastly, a significant three-way interaction was found ($F = 2.649, p = .00821$), indicating that the combined effect of noise types, noise levels, and control conditions on perceived loudness is complex, with each factor influencing the others in a nonlinear way. Overall, the analysis revealed that all three factors—noise types, noise control conditions, and noise levels—have significant effects on perceived loudness, with complex interactions between them. The results emphasize the importance of considering these factors together when studying noise perception and developing noise control strategies. Results of the 3WR-ANOVA are findable in Table 6.

The post hoc analysis using the Tukey HSD method reveals significant differences in the study of perceived loudness (PLN) between groups based on noise level, noise type, and control conditions. Regarding noise type comparisons, a significant difference was observed between motorbike and street ($p = .021, d = -2.917$), suggesting distinct PLN levels between these two types. A highly significant difference was also found between motorbike and train ($p = .00012, d = -4.6603$), indicating a pronounced distinction in perception. However, no significant difference was detected between street and train ($p = 0.24, d = -1.743$), reflecting a relative similarity in the perceived loudness. For noise control conditions, the SFANC-NLMS algorithm demonstrated a highly significant reduction in PLN compared to No Control ($p < 0.0001, d = -8.473$), showcasing its strong efficacy. Additionally, SFANC-NLMS outperformed NLMS with a significant improvement ($p = 2.37 \times 10^{-6}, d = -5.1047$). While NLMS also showed a significant reduction in PLN compared to No Control ($p = 0.00179, d = -3.3682$), it remains less effective than SFANC-NLMS. Significant differences were also observed across all pairwise comparisons of noise levels (55 dB, 65 dB, and 72 dB), with $p < 0.0001$. The effect sizes were -11.968 for 55 dB vs. 65 dB, -22.419 for 55 dB vs. 72 dB, and -10.451 for 65 dB vs. 72 dB. These findings confirm that PLN increases systematically with rising noise levels. When considering interactions between noise type, level, and control conditions, highly significant differences were observed ($p < 0.0001$) with effect sizes reaching up to -35.086. This demonstrates the cumulative influence of these factors on PLN. In extreme conditions, such as comparing SFANC-NLMS (Moto, 55 dB) to No Control (Train, 72 dB), the differences were strikingly significant ($p = 1.77 \cdot 10^{-11}, d = -35.086$), emphasizing the substantial impact of the SFANC-NLMS algorithm, particularly when noise type and level disparities are most pronounced. Overall, the results highlight the superior performance of the SFANC-NLMS algorithm compared to NLMS and No control conditions, especially under challenging high-noise scenarios. The findings also underscore the critical role of noise level and type in shaping PLN, further

Table 7
Post hoc Tukey HSD results for significant comparisons in the PLN study.

Comparison	p-value	Effect
Moto vs. Street	0.021 *	-2.92
Moto vs. Train	0.00012 ***	-4.66
Street vs. Train	0.24	-1.74285
SFANC-NLMS vs. No Control	< 0.0001 ****	-8.47
SFANC-NLMS vs. NLMS	< 0.0001 ****	-5.11
No Control vs. NLMS	0.002 **	3.37
55 dB vs. 65 dB	< 0.0001 ****	-11.97
55 dB vs. 72 dB	< 0.0001 ****	-22.42
65 dB vs. 72 dB	< 0.0001 ****	-10.45
Moto (55 dB) vs. Moto (72 dB)	< 0.0001 ****	-22.66
Moto (55 dB) vs. Street (72 dB)	< 0.0001 ****	-24.11
Moto (55 dB) vs. Train (72 dB)	< 0.0001 ****	-29.54
Street (55 dB) vs. Train (72 dB)	< 0.0001 ****	-22.77
Train (55 dB) vs. Moto (72 dB)	< 0.0001 ****	-20.37
Train (55 dB) vs. Street (72 dB)	< 0.0001 ****	-21.83
Train (55 dB) vs. Train (72 dB)	< 0.0001 ****	-27.26
SFANC-NLMS (55 dB) vs. SFANC-NLMS (72 dB)	< 0.0001 ****	-20.88
SFANC-NLMS (55 dB) vs. No Control (72 dB)	< 0.0001 ****	-30.77
SFANC-NLMS (55 dB) vs. NLMS (72 dB)	< 0.0001 ****	-25.92
No Control (55 dB) vs. No Control (72 dB)	< 0.0001 ****	-24.08
NLMS (55 dB) vs. No Control (72 dB)	< 0.0001 ****	-27.15
NLMS (55 dB) vs. NLMS (72 dB)	< 0.0001 ****	-22.31
SFANC-NLMS (55 dB) vs. No Control (72 dB)	< 0.0001 ****	-20.52
SFANC-NLMS (Moto, 55 dB) vs. No Control (Street, 72 dB)	< 0.0001 ****	-32.89
SFANC-NLMS (Moto, 55 dB) vs. No Control (Train, 72 dB)	< 0.0001 ****	-35.09
SFANC-NLMS (Moto, 55 dB) vs. NLMS (Train, 72 dB)	< 0.0001 ****	-33.91
No Control (Moto, 55 dB) vs. No Control (Train, 72 dB)	< 0.0001 ****	-30.43
NLMS (Moto, 55 dB) vs. No Control (Street, 72 dB)	< 0.0001 ****	-30.86
NLMS (Moto, 55 dB) vs. No Control (Train, 72 dB)	< 0.0001 ****	-33.06
NLMS (Moto, 55 dB) vs. NLMS (Train, 72 dB)	< 0.0001 ****	-31.89
SFANC-NLMS (Street, 55 dB) vs. No Control (Train, 72 dB)	< 0.0001 ****	-32.66
SFANC-NLMS (Street, 55 dB) vs. No Control (Train, 72 dB)	< 0.0001 ****	-29.66

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

validating the effectiveness of advanced noise control techniques. Results of the post hoc are findable in Table 7.

5. Discussion

The following discussion seeks to address the research questions established in Section 1. Section 5.1 examines the impact of ANC on perceived annoyance under different conditions (RQ1). Section 5.2 evaluates the influence of noise level on perceived annoyance under different ANC conditions (RQ2). Finally, Section 5.3 explores the interaction between noise type, ANC condition, and noise level (RQ3).

5.1. Effect of ANC algorithms on perceived annoyance

The results of this study provide insights into how different ANC algorithms (NLMS and SFANC-NLMS) influence perceived annoyance in a simulated vehicle environment, addressing RQ1. Overall, the SFANC-NLMS algorithm was more effective in reducing perceived annoyance compared to the traditional NLMS algorithm, across almost all noise types and levels. This was particularly evident for train and motorbike noises, where SFANC-NLMS significantly reduced perceived annoyance. For motorbike noise, SFANC-NLMS also outperformed NLMS, highlighting its ability to manage more complex noise characteristics. While both algorithms showed similar performance in some cases, SFANC-NLMS consistently resulted in lower annoyance scores, suggesting that it is better suited for handling various noise types. Noise level also played a crucial role, with higher levels (72 dB) leading to greater perceived annoyance, regardless of the ANC condition. Interestingly, the interaction between noise type and ANC algorithm was not significant, indicating that the effects of the algorithms and noise types were more influential individually. As shown in Fig. 6, stimuli controlled by SFANC-NLMS generally resulted in the lowest perceived annoyance scores (PAY), confirming its superior effectiveness.

5.2. Effect of sound level on perceived annoyance under different ANC conditions

The results regarding perceived annoyance also reveal that sound level (55, 65, and 72 dB(A)) moderates the relationship between ANC algorithm type and perceived annoyance, addressing RQ2. In general, higher sound levels were associated with higher perceived annoyance scores, even with the application of ANC. This emphasizes the importance of controlling sound intensity to reduce annoyance. The moderating effect of sound level seemed most pronounced at 72 dB, where the differences between algorithms were most evident. However, the interaction between ANC algorithm and sound level did not significantly alter the direction of the relationship between these variables and perceived annoyance. In summary, sound level affects annoyance, but its impact remains consistent across different ANC algorithms.

5.3. Interaction between noise sources, ANC condition, and sound levels

Regarding the interaction between noise sources, ANC algorithms, and sound levels, the results show that different noise sources (e.g., motorcycle, street, train noises) interact with ANC algorithms and sound levels to shape perceived annoyance, addressing RQ3. In particular, street and train noises showed a stronger interaction with ANC algorithms, with SFANC-NLMS significantly reducing perceived annoyance compared to NLMS. For motorcycle noise, while SFANC-NLMS also showed greater reductions in annoyance, the interaction between noise type and ANC algorithm was not significant in all cases. This suggests that, while ANC algorithms play a key role in managing different noise

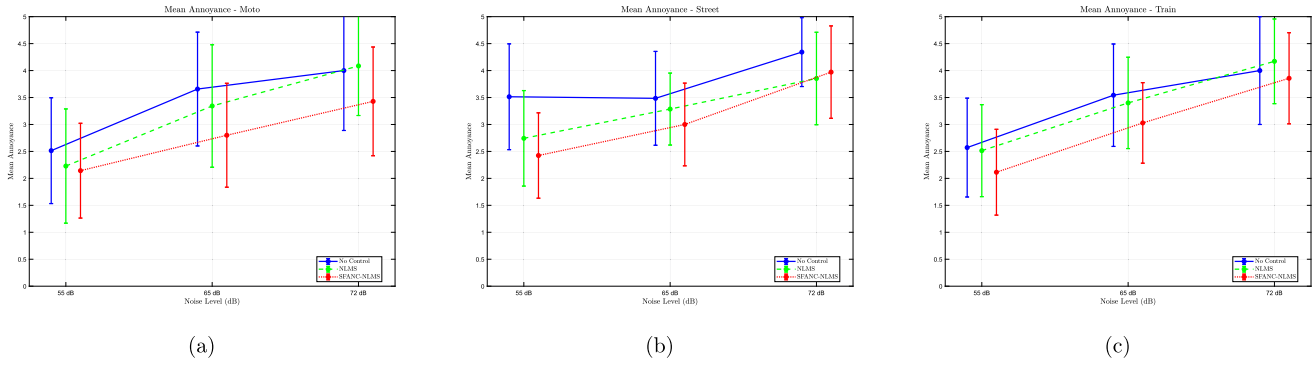


Fig. 6. Means and standard deviations of the PAY score obtained for different noise levels for (a) Motorbike noise, (b) Street noise, and (c) Train noise.

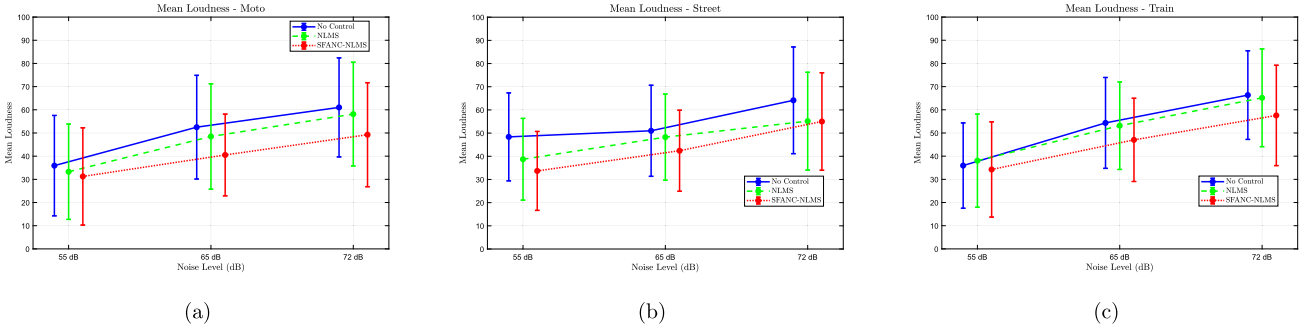


Fig. 7. Means and standard deviations of the PLN score obtained for different noise levels for (a) Motorbike noise, (b) Street noise, and (c) Train noise.

sources, the spectral characteristics of the noise have a stronger influence on the algorithm's effect on perceived annoyance. Sound level also had a significant impact, with higher noise levels leading to greater perceived annoyance, regardless of the ANC algorithm or noise type. Additionally, Fig. 7 provides a visualization of the variation in perceived loudness (PLN) scores across different noise types and sound levels, complementing our findings on annoyance perception. As shown in the figure, stimuli controlled by SFANC-NLMS generally resulted in the lowest PLN mean scores, further supporting the conclusion that SFANC-NLMS provides a more comfortable acoustic environment by reducing both annoyance and loudness.

6. Conclusion

This study assessed the impact of different ANC algorithms on perceived annoyance, loudness, and affective quality.

RQ1: To what extent do different ANC algorithms influence perceived annoyance? The results show that the SFANC-NLMS algorithm is more effective than the traditional NLMS algorithm in reducing perceived annoyance, particularly for complex noise sources such as motorbike and train noises.

RQ2: Does sound level moderate the relationship between ANC algorithm and perceived annoyance? Sound level significantly influences perceived annoyance, with higher noise levels (72 dB) leading to greater annoyance, regardless of the ANC condition. However, the effect of sound level did not alter the relationship between ANC algorithm type and perceived annoyance.

RQ3: How do different noise sources interact with ANC algorithms and sound levels to shape perceived annoyance? Different noise sources interacted with ANC algorithms and sound levels, with SFANC-NLMS consistently reducing perceived annoyance more effectively than NLMS across all noise types. Street and train noises showed stronger interactions with ANC algorithms, while the motorbike noise exhibited more complexity.

In conclusion, while objective metrics have already demonstrated the superior performance of the SFANC-NLMS algorithm, this study confirms its effectiveness from a subjective perspective, showing its ability to reduce perceived annoyance and loudness and enhance acoustic comfort. However, noise level remains a key factor influencing subjective experience, emphasizing the need for sound intensity control. Further research is required to optimize SFANC-NLMS for more complex noise profiles and assess its applicability in real-world conditions.

CRediT authorship contribution statement

Alkahf Aboutiman: Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Project administration, Methodology, Formal analysis, Data curation, Conceptualization. **Zulfi Rachman:** Writing – review & editing, Validation, Resources, Formal analysis, Data curation, Conceptualization. **Tin Oberman:** Writing – review & editing, Validation, Supervision, Resources, Formal analysis, Data curation, Conceptualization. **Francesco Alletta:** Writing – review & editing, Validation, Supervision, Resources, Formal analysis, Data curation, Conceptualization. **Jian Kang:** Writing – review & editing, Validation, Supervision, Conceptualization. **Hamid Reza Karimi:** Writing – review & editing, Validation, Supervision, Conceptualization. **Francesco Ripamonti:** Writing – review & editing, Validation, Supervision, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

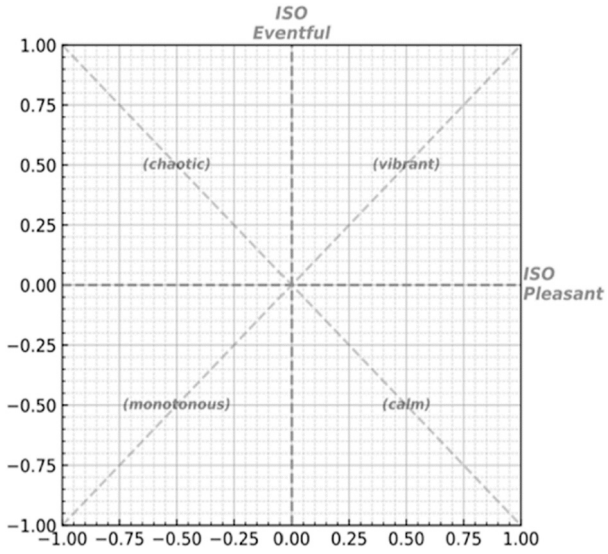
The work was supported by the project IN-NOVA: Active reduction of noise transmitted into and from enclosures through encapsulated structures, which has received funding from the European Union's Horizon Europe Programme under the Marie Sklodowska-Curie grant agreement no. 101073037.

Appendix A. Questionnaires

Table A.1
Pre-test assessment participant information questionnaire.

Question Category	Specific Questions	Rating Scale
Individual Noise Sensitivity (INS)	I am sensitive to noise. I find it difficult to relax in a place that's noisy. I get mad at people who make noise that keeps me from falling asleep or getting work done. I get annoyed when my neighbors are noisy. I get used to most noises without much difficulty.	Totally Disagree–Totally Agree (10-point categorical)
Baseline Noise Annoyance (BNA)	When you are at home, how much does noise from road traffic bother, disturb or annoy you? When you are at home, how much does noise from aircraft bother, disturb or annoy you? When you are at home, how much does noise from train bother, disturb or annoy you? When you are at home, how much does noise from children bother, disturb or annoy you? When you are at home, how much does noise from other people bother, disturb or annoy you? When you are at home, how much does noise from animals bother, disturb or annoy you?	Not at all–Extremely (5-point categorical)
Perceived Stress Scale (PSS-4)	How often have you felt that you were unable to control the important things in your life? How often have you felt confident about your ability to handle your personal problems? How often have you felt that things were going your way? How often have you felt difficulties were piling up so high that you could not overcome them?	Never–Very often (5-point categorical)

Table A.2
Noise Evaluation Questionnaire: PAY, PAQ, and PLN.

Question Category	Specific Questions	Rating Scale
Perceived Annoyance (PAY)	Thinking about the noise you just heard, how much does the noise bother, disturb, or annoy you?	Not at all–Extremely (5-point categorical)
Perceived Affective Quality (PAQ)	Thinking about the sound you just heard, where would you place the surrounding sound environment on this scale?	
		
Perceived Loudness (PLN)	How loud would you say the sound environment is?	Scale between 0–100

Data availability

Data will be made available on request.

References

[1] Cabell RH, Fuller C. A principal component algorithm for feedforward active noise and vibration control. J Sound Vib 1999;227(1):159–81.
[2] Goodwin GC, Silva EI, Quevedo DE. Analysis and design of networked control systems using the additive noise model methodology. Asian J Control 2010;12(4):443–59.
[3] Hartmann WM. Signals, sound, and sensation. Springer Science & Business Media; 2004.
[4] Li Y, Wang X, Huang R, Qiu Z. Active vibration and noise control of vibro-acoustic system by using pid controller. J Sound Vib 2015;348:57–70.
[5] Wilby JF. Aircraft interior noise. J Sound Vib 1996;190(3):545–64.
[6] Misol M, Algermissen S, Monner HP. Experimental investigation of different active noise control concepts applied to a passenger car equipped with an active windshield. J Sound Vib 2012;331(10):2209–19.
[7] Maamoun KSA, Wrona S, Pawelczyk M, Karimi HR. Shaping of the frequency response of vibrating plates with openings for vibro-acoustic systems. Mech Syst Signal Process 2024;218:111539.
[8] Fuller C, Hansen C, Snyder S. Active control of sound radiation from a vibrating rectangular panel by sound sources and vibration inputs: an experimental comparison. J Sound Vib 1991;145(2):195–215.
[9] Li Y, Cheng L. Active noise control of a mechanically linked double panel system coupled with an acoustic enclosure. J Sound Vib 2006;297(3–5):1068–74.
[10] Shorter P. Recent advances in automotive interior noise prediction. In: SAE BRASIL noise and vibration conference; 2008.
[11] Rustighi E, Elliott S. Stochastic road excitation and control feasibility in a 2d linear tyre model. J Sound Vib 2007;300(3–5):490–501.
[12] Belgacem W, Berry A, Masson P. Active vibration control on a quarter-car for cancellation of road noise disturbance. J Sound Vib 2012;331(14):3240–54.

[13] Ranjan R, Gan W-S. Natural listening over headphones in augmented reality using adaptive filtering techniques. IEEE/ACM Trans Audio Speech Lang Process 2015;23(11):1988–2002.
[14] Yang F, Cao Y, Wu M, Albu F, Yang J. Frequency-domain filtered-x lms algorithms for active noise control: a review and new insights. Appl Sci 2018;8(11):2313.
[15] Reddy RM, Panahi IM, Briggs R. Hybrid fxrls-fxnlms adaptive algorithm for active noise control in fmri application. IEEE Trans Control Syst Technol 2010;19(2):474–80.
[16] Guo J, Yang F, Yang J. Convergence analysis of the conventional filtered-x affine projection algorithm for active noise control. Signal Process 2020;170:107437.
[17] Cha Y-J, Choi W, Suh G, Mahmoudkhani S, Büyüköztürk O. Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types. Comput-Aided Civ Infrastruct Eng 2018;33(9):731–47.
[18] Deng L, Yu D, et al. Deep learning: methods and applications. Found Trends Signal Process 2014;7(3–4):197–387.
[19] Zhang H, Wang D. Deep anc: a deep learning approach to active noise control. Neural Netw 2021;141:1–10.
[20] Cha Y-J, Mostafavi A, Benipal SS. Dnoisenet: deep learning-based feedback active noise control in various noisy environments. Eng Appl Artif Intell 2023;121:105971.
[21] Luo Z, Shi D, Gan W-S. A hybrid sfanc-fxnlms algorithm for active noise control based on deep learning. IEEE Signal Process Lett 2022;29:1102–6.
[22] Belyi V, Gan W-S. Integrated psychoacoustic active noise control and masking. Appl Acoust 2019;145:339–48.
[23] Gonzalez A, Ferrer M, De Diego M, Pinero G, Garcia-Bonito J. Sound quality of low-frequency and car engine noises after active noise control. J Sound Vib 2003;265(3):663–79.
[24] Fastl H. The psychoacoustics of sound-quality evaluation. Acta Acust Acust 1997;83(5):754–64.
[25] Guski R. Psychological methods for evaluating sound quality and assessing acoustic information. Acta Acust Acust 1997;83(5):765–74.
[26] Van der Auweraer H, Wyckaert K, Hendricx W. From sound quality to the engineering of solutions for nvh problems: case studies. Acta Acust Acust 1997;83(5):796–804.
[27] Bisping R. Car interior sound quality: experimental analysis by synthesis. Acta Acust Acust 1997;83(5):813–8.
[28] Camacho A, Pinero G, De Diego M, Gonzalez A. Exploring roughness perception in car engine noises through complex cepstrum analysis. Acta Acust Acust 2008;94(1):130–40.
[29] Brandl FK, Biermayer W. A new tool for the onboard objective assessment of vehicle interior noise quality. Tech. Rep., SAE Technical Paper. 1999.
[30] Miśkiewicz A, Letowski T. Psychoacoustics in the automotive industry. Acta Acust Acust 1999;85(5):646–9.
[31] Duan J. Active control of vehicle powertrain and road noise. University of Cincinnati; 2011.
[32] Ryu S, Lee Y-S, Kim S. Active control of engine sound quality in a passenger car using a virtual error microphone. Int J Parallel Program 2020;48(5):909–27.
[33] Chen W, Lu C, Liu Z, Li X, Shen J. A novel feedforward hybrid active sound quality control algorithm for both narrowband and broadband sound-profiling. Appl Acoust 2021;180:108095.
[34] Wang X, Song Y, Wang T, Wang Y, Liu N. Hybrid vibro-acoustic active control method for vehicle interior sound quality under high-speed. Appl Acoust 2022;186:108419.
[35] Schiffbänker H, Brandl F, Thien G. Development and application of an evaluation technique to assess the subjective character of engine noise. SAE Transact 1991:1536–46.
[36] Ingham R, Otto N, McCollum T. Sound quality metric for diesel engines. SAE Transact 1999:3003–7.
[37] Hussain M, Gölles J, Ronacher A, Schiffbänker H. Statistical evaluation of an annoyance index for engine noise recordings. SAE Transact 1991:1527–35.
[38] Khan MS, Johansson Ö, Sundbäck U. Development of an annoyance index for heavy-duty diesel engine noise using multivariate analysis. Noise Control Eng J 1997;45(4):157–67.
[39] Liu L. Design, simulation and testing of a coupled plate-cavity system targeted for vehicle interior noise analysis and control. PhD thesis. Politecnico di Milano; 2021.
[40] Wu L, Qiu X, Guo Y. A generalized leaky fxlms algorithm for tuning the waterbed effect of feedback active noise control systems. Mech Syst Signal Process 2018;106:13–23.
[41] Milani AA, Kannan G, Panahi IM. On maximum achievable noise reduction in anc systems. In: 2010 IEEE international conference on acoustics, speech and signal processing. IEEE; 2010. p. 349–52.
[42] Kuo SM, Morgan DR. Active noise control systems, vol. 4. New York: Wiley; 1996.
[43] Aboutiman A. A convolutional neural network-based hybrid active noise control in encapsulated structures with non-minimum phase characteristics. J Acoust Soc Am 2025. submitted for publication.
[44] Cartwright M, Cramer J, Mendez AEM, Wang Y, Wu H-H, Lostanlen V, et al. Sonycust-v2: an urban sound tagging dataset with spatiotemporal context. arXiv preprint. arXiv:2009.05188, 2020.
[45] Aletta F, Vander Mynsbrugge T, Van de Velde D, De Vriendt P, Thomas P, Filipan K, et al. Awareness of 'sound'in nursing homes: a large-scale soundscape survey in Flanders (Belgium). Build Acoust 2018;25(1):43–59.
[46] I. O. for Standardization, Iso/ts 15666 acoustics — assessment of noise annoyance by means of social and socio-acoustic surveys (2nd ed.). Technical Specification ISO/TS 15666. International Organization for Standardization, Geneva, Switzerland, 2021.
[47] Lam B, Ong Z-T, Ooi K, Ong W-H, Wong T, Watcharasupat KN, et al. Automating urban soundscape enhancements with ai: in-situ assessment of quality and restorativeness in traffic-exposed residential areas. Build Environ 2024;266:112106.
[48] Karam F, Bérard A, Sheehy O, Huneau M-C, Briggs G, Chambers C, et al. Reliability and validity of the 4-item perceived stress scale among pregnant women: results from the otis antidepressants study. Res Nurs Health 2012;35(4):363–75.
[49] Mitchell A, Oberman T, Aletta F, Erfanian M, Kachlicka M, Lionello M, et al. The soundscape indices (ssid) protocol: a method for urban soundscape surveys—questionnaires with acoustical and contextual information. Appl Sci 2020;10(7):2397.
[50] Lam B, Lim KCQ, Ooi K, Ong Z-T, Shi D, Gan W-S. Anti-noise window: subjective perception of active noise reduction and effect of informational masking. Sustain Cities Soc 2023;97:104763.
[51] I.O. for Standardization. Acoustics — soundscape — part 2: data collection and reporting requirements. Tech. Rep., ISO/TS, 2018. ISO/TS 12913-2. 2018.
[52] ISO. Iso 532-1. Acoustics—methods for calculating loudness—part 1: Zwicker method. 2017.
[53] Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). Biometrika 1965;52(3–4):591–611.
[54] Levene H. Robust tests for equality of variances. 1960.
[55] Mauchly JW. Significance test for sphericity of a normal n-variate distribution. Ann Math Stat 1940;11(2):204–9.
[56] Mardia KV. Measures of multivariate skewness and kurtosis with applications. Biometrika 1970;57(3):519–30.

Expert Systems With Applications 309 (2026) 131247

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Hybrid deep learning–based active noise control for encapsulated structures with openings

Alkahf Aboutiman [a,*], Khaled Said Ahmed Maamoun [b], Hamid Reza Karimi [a], Francesco Ripamonti [a]

[a] Department of Mechanical Engineering, Politecnico di Milano, Via La Masa 1, 20156, Milan, Italy
[b] Department of Measurements and Control Systems, Silesian University of Technology, Akademicka 2A, 44–100, Gliwice, Poland

ARTICLE INFO

Keywords:
Active noise control
Learning-based weighting
Encapsulated structure with opening
Loudspeaker nonlinearity

ABSTRACT

Traditional active noise control (ANC) systems rely on adaptive filtering techniques that assume a linear relationship between signals. However, in practical implementations, nonlinear distortions, particularly those introduced by loudspeakers, can degrade control performance, limiting the effectiveness of conventional ANC approaches. This paper presents a hybrid ANC algorithm, GFANC-THFxNLMS, designed to enhance adaptability and robustness in complex acoustic environments, such as encapsulated structures with openings. The proposed method leverages deep learning to construct a dataset of pre-trained sub-control filters and employs a 1D convolutional neural network (CNN) to dynamically generate the most suitable filter for the incoming noise. An adaptive filtering stage then fine-tunes the control filter in real time, ensuring optimal noise attenuation. The performance of GFANC-THFxNLMS is evaluated against conventional ANC algorithms, namely GFANC and THFxNLMS, under both time-invariant and time-varying conditions. The results indicate that the proposed method achieves enhanced noise reduction, with 6.74 dB using THFxNLMS, 8.01 dB using GFANC, and 9.49 dB using GFANC-THFxNLMS, highlighting its effectiveness in noise mitigation.

1. Introduction

Exposure to excessive noise is recognized as a serious threat to human well-being, with significant impacts on comfort, health, and safety (World Health Organization, 2018). The advancement of industrial machinery and domestic devices has contributed to elevated sound levels, potentially compromising both user comfort and operational safety. Encapsulated structures are commonly used as noise barriers to isolate sound sources from their surroundings (Isaac et al., 2022; Wrona et al., 2022). However, incorporating openings in these structures-necessary for airflow, heat dissipation, and device monitoring compromises their ability to reduce noise transmission (Maamoun et al., 2024, 2025). This trade-off is essential in many engineering and industrial applications. It has increased the demand for advanced noise reduction techniques, such as ANC. The principle of superposition serves as the foundation for ANC, a technique that reduces unwanted noise by generating an anti-noise signal of the same amplitude but opposite phase, thus canceling the disturbance (Goodwin et al., 2010; Jeong & Park, 2023). Unlike passive noise control, which relies on sound-absorbing materials, ANC actively adapts in real time by manipulating acoustic signals (Aboutiman et al., 2025a; Lam et al., 2020; Tan & Wang, 2019). This real-time adaptability makes ANC a valuable solution in fields such as domestic appliances, construction, and manufacturing, where passive methods are limited by cost and space constraints (Hartmann, 2004).

In enclosed environments, passive noise control remains the predominant approach, primarily relying on materials engineered for sound absorption. However, such methods become less effective at low frequencies due to the substantial mass required for adequate performance (Aboutiman et al., 2025b; Fuller et al., 1991; Misol et al., 2012). This limitation is particularly significant for many domestic appliances and industrial machines, where the dominant noise components typically lie below 500 Hz. One implementation method for ANC in enclosures with openings, known as boundary control (Wang et al., 2017), involves placing secondary actuators along the perimeter of the opening. This configuration improves airflow since the speakers do not obstruct the aperture. However, the effectiveness of this method may be limited by physical constraints imposed by the boundary length. It also requires a boundary thick enough to mount the actuators securely. In contrast, planar control arranges the secondary actuators in a grid pattern across the plane of the opening (Murao et al., 2016; Tao et al., 2016). Although

* Corresponding author.
E-mail addresses: alkahfahmat.aboutiman@polimi.it (A. Aboutiman), khaled.maamoun@polsl.pl (K.S.A. Maamoun), hamidreza.karimi@polimi.it (H.R. Karimi), francesco.ripamonti@polimi.it (F. Ripamonti).

https://doi.org/10.1016/j.eswa.2026.131247
Received 27 June 2025; Received in revised form 14 January 2026; Accepted 14 January 2026
Available online 16 January 2026
0957-4174/© 2026 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

this approach eliminates the need for thick boundaries, it reduces access to the noise source because the actuator grid partially blocks the opening. However, none of these approaches discuss the effects of nonlinearities in such systems.

Nonlinearities in ANC systems present significant challenges, as they can degrade the performance of adaptive filtering techniques. Traditional ANC relies on adaptive filters, such as the widely used Filtered-x Least Mean Square (FxLMS) algorithm, which estimates the secondary path and minimizes an error signal to optimize noise cancellation (Elliott et al., 1987; Manolakis et al., 2005). However, real-world ANC applications introduce nonlinear distortions due to hardware imperfections in amplifiers and loudspeakers, which result in inaccurate secondary path estimation and reduced noise attenuation (Kuo & Morgan, 1996). Several approaches have been proposed to address these issues. The Volterra expansion has proven effective in modeling weak nonlinearities, and second-order Volterra-based FxLMS algorithms have been developed to improve feedforward ANC performance (Guo et al., 2018; Lashkari, 2006). Additionally, nonlinear models such as the Nonlinear AutoRegressive model with eXogenous variables (NARX) have been employed to enhance ANC controller efficiency (Napoli & Piroddi, 2009). To mitigate saturation-type nonlinearities, variants like the Tangential Hyperbolic Function FxLMS (THF-FxLMS) and bilinear FxLMS have been proposed (Ghasemi et al., 2016; Kuo & Wu, 2005). More recent methods, including FXlogLMS, FXatanLMS, and FXsigLMS, use nonlinear transformations such as logarithmic, inverse tangent, and sigmoid functions. These techniques help avoid threshold estimation and improve convergence (Jun et al., 2012; Wu et al., 2010; Zhou et al., 2015). While such advancements improve control accuracy, they often increase computational cost. To balance efficiency and robustness, hybrid algorithms like MGFxRLS-NSSFxLMS and GMACFxLMS have been introduced (Meng & Chen, 2020). In parallel, feedback ANC systems, which rely solely on error microphones to measure residual noise, have been explored as alternatives to feedforward ANC in reverberant environments where reference signals are distorted (Lam et al., 2021; Zhang & Wang, 2021). However, traditional feedback systems struggle with broadband noise attenuation and require precise controller design to maintain stability. Recent advances in robust control, including fuzzy-based and fault-tolerant strategies, show potential for nonlinear stochastic systems but remain underexplored in ANC applications (Liang et al., 2022; Sun et al., 2020). Despite significant progress, efficiently mitigating nonlinear distortions in ANC remains an active area of research.

To address these limitations, neural networks have been explored for nonlinear ANC due to their ability to model complex relationships. Early studies introduced multilayer perceptron (MLP) networks for active vibration control, where adaptive filtered-x backpropagation was used to update network weights (Snyder & Tanaka, 1995). Further advancements led to improved training algorithms that enhanced convergence speed while reducing computational cost (Bouchard et al., 1999; Chang & Luoh, 2007). Alternative architectures, such as functional link neural networks (Das & Panda, 2004; Krukowicz, 2010), radial basis function networks (Tokhi & Wood, 1997), fuzzy neural networks (Zhang et al., 2006), and recurrent neural networks (Bambang, 2008), have also been employed to mitigate nonlinear distortions. These approaches demonstrated the potential of artificial intelligence in ANC by using online adaptation to continuously adjust control parameters. With the advent of deep learning, new opportunities have emerged for tackling nonlinear ANC problems. Deep neural networks (DNNs) can model highly nonlinear and time-varying systems, allowing them to adapt to complex acoustic environments. Unlike traditional machine learning models that require manual feature extraction, deep learning methods-such as convolutional neural networks (CNNs) and long short-term memory (LSTM) networks-can automatically learn spatial and temporal features relevant to noise control. CNN-based ANC models have shown superior performance by leveraging hierarchical feature extraction, which improves noise reduction across various environments (Cha et al., 2018). Meanwhile, LSTM networks excel at capturing long-term temporal dependencies, making them particularly effective for predicting and canceling nonstationary noise patterns (Duan et al., 2016; Fu et al., 2016). Recent studies have demonstrated that deep learning-based ANC approaches can outperform conventional adaptive filtering methods. For instance, convolutional recurrent networks combining CNN and LSTM architectures have been proposed as alternatives to Volterra-based FxLMS controllers and MLP-based ANC systems. These deep models show improved generalization across noise types and environments, overcoming the limitations of single-filter adaptive methods (Cha et al., 2023; Guo et al., 2018; Zhang & Wang, 2021). In parallel, hybrid learning-based architectures have been explored in real-time embedded applications under strict resource constraints, such as threat detection in IoT networks. These studies show that combining lightweight neural networks for feature reduction with classical adaptive stages enables efficient and robust real-time processing, which is well aligned with the requirements of embedded ANC systems (Elzaghmouri et al., 2024; Sivasakthi et al., 2024).

This paper introduces a hybrid ANC algorithm, GFANC-THFxNLMS, which leverages deep learning to address the nonlinearities introduced by loudspeakers in ANC applications for encapsulated structures with openings. The proposed method builds upon the GFANC framework described in (Luo et al., 2025, 2023), which enhances traditional ANC algorithms such as SFANC by decomposing a pre-trained broadband control filter into multiple sub-filters (Vetterli & Le Gall, 1989). A lightweight 1D CNN running on a co-processor then determines the optimal combination weights based on the characteristics of the primary noise.This design minimizes computational load and simplifies the ANC setup, as it relies on a single pre-trained filter rather than multiple independently trained ones. In addition, the GFANC-THFxNLMS hybrid improves performance by adaptively updating the control filter coefficients, following the adaptive strategy introduced in (Luo et al., 2022). The main contributions of this paper are threefold: (i) developing an ANC algorithm that uses deep learning to overcome the limitations caused by nonlinearities; (ii) implementing and validating this algorithm in a complex acoustic environment, specifically an encapsulated structure with an opening; and (iii) proposing a 1D CNN architecture that achieves high accuracy while keeping the number of parameters minimal.

The rest of this paper is structured as follows: The paper first presents the experimental setup used to compute the impulse response of the control environment. Secondly, it introduces the ANC algorithm along with its neural network architecture and training process. The results obtained with the proposed algorithm are then presented and compared with traditional algorithms. Finally, the paper concludes the study and highlights its contributions.

2. System identification

The identification experiments were conducted in a laboratory room measuring 5.8 m in length and 3.5 m in width. While the walls are covered with sound-absorbing materials, the presence of various equipment and objects results in acoustic conditions that more closely resemble a typical operational environment than those of an anechoic chamber.
The experimental setup consists of a rectangular steel plate mounted on a concrete enclosure that houses the primary noise source, as shown in Fig. 1. A Behringer EUROLIVE B208D speaker is employed as the primary excitation source, generating band-limited random noise with frequencies up to 1500 Hz with a maximum sound pressure level (SPL) of 113 dB. The secondary source is a loudspeaker (JBL Stage2 424) located inside the enclosure, positioned centrally behind the opening. The Analog Device SC598 with SHARC+ core is used along with EVSOMCRR-EZKIT development board (Analog Devices, 2024). The plate is made of steel, characterized by a density of $\rho = 7850\,\mathrm{kg/m^3}$ and a Young's modulus of $E = 210 \times 10^9$ Pa. Its dimensions are $a = 0.42$ m (length), $b = 0.39$ m (width), and $h = 0.001$ m (thickness). This setup, with the given parameters, is used to identify the primary and secondary paths using Mic 1, as it represents a wide range of engineering applications involving encapsulated structures with openings.

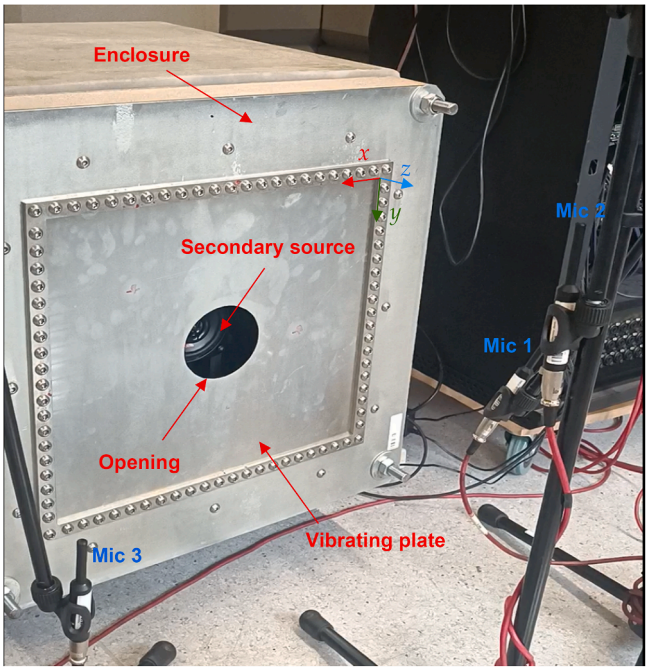


Fig. 1. Overview of the experimental rig showing a vibrating plate with a central circular opening of radius 0.05 m. The steel plate spans 0.42 m in x, 0.39 m in y, and has a thickness of 0.001 m in $-z$. Microphones (Mic 1, Mic 2, and Mic 3) are positioned at coordinates (-0.18 m, 0.12 m, 0.25 m), (0.21 m, -0.1 m, 0.4 m), and (0.53 m, 0.24 m, 0.25 m), respectively. Mic 1 is used as the error microphone, while Mic 2 and Mic 3 are intended for observation, if needed. The primary noise source is located at (0.3 m, 0.1 m, 0.1 m) and the loudspeaker, used as a secondary source, is centered behind the opening.

For path identification, the Normalized Least Mean Square (NLMS) algorithm is employed. The original signals are recorded at a sampling rate of 48 kHz and subsequently downsampled to 3 kHz to reduce computational complexity, given that the frequency content of interest lies below the Nyquist limit. A filter length of 300 coefficients is used for both the primary and secondary paths. The corresponding transfer functions, obtained from the measured impulse responses, are presented in Fig. 2. The primary path exhibits prominent resonant behavior, particularly below 350 Hz, due to the complex noise barrier characteristics influenced by both the radiated sound from the vibrating plate and the sound leakage through the opening (Maamoun et al., 2024). In contrast, the secondary path, which originates from a secondary loudspeaker positioned near an opening, is less affected by structural vibrations. Its response is more attenuated and smoother, particularly at mid-to-high frequencies, indicating a more direct acoustic path with minimal structural coupling. The phase responses of both paths exhibit smooth behavior which indicates physically consistent acoustic paths' modeling , as in Fig. 2.

3. The GFANC-THFxNLMS algorithm

This section presents the GFANC-THFxNLMS based on deep learning and the configuration used to make it effective in the environment presented in the previous section. The block diagram of the algorithm is presented in Fig. 3.

3.1. The tangent hyperbolic filtered-x normalized least mean square algorithm description

As presented in (Ghasemi et al., 2016), the Tangent Hyperbolic Filtered-x Normalized Least Mean Square (THFxNLMS) algorithm is an adaptive algorithm designed to enhance the robustness of ANC systems, particularly in the presence of nonlinearities. This section presents the THFxNLMS algorithm, which extends the standard FxNLMS method by incorporating a hyperbolic tangent transformation to improve stability and performance. The block diagram of the THFxNLMS algorithm is the one presented in Fig. 3 without the co-processor, where $P(z)$ represents the transfer function of the primary path, $W(z)$ denotes the control filter, $S(z)$ is the transfer function of the secondary path, and its estimate is denoted as $\hat{S}(z)$.

The reference signal vector $x(n)$ is defined as $x(n) = [x(n), x(n-1), \ldots, x(n-L+1)]^T$. Let T denote the transpose operation and L denotes the length of the control filter. The resulting output of the control filter is therefore $y(n)$, as defined in Eq. (1).

$$y(n) = w^T(n)x(n), \tag{1}$$

Where $w(n)$ represents the coefficient vector of the control filter. The residual error signal can be expressed as follows:

$$e(n) = d(n) - s(n) * \left(w^T(n)x(n)\right), \tag{2}$$

where $d(n)$ and $s(n)$ denote the primary disturbance and the impulse response of the secondary path, respectively, and $*$ is the linear convolution. A key difference introduced by the THFxNLMS algorithm is the use of a nonlinear transformation for the residual error signal $e(n)$. The transformation function is defined as:

$$f(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \tag{3}$$

This function is odd-symmetric, continuous, monotonic, and bounded within the range $(-1, 1)$. By applying this transformation, the variance of the error signal remains finite, improving the stability of the learning process and making the algorithm more resilient to nonlinear distortions. The loss function of the THFxNLMS algorithm is given by:

$$J(n) = \frac{f^2(b(n)e(n))}{2 \cdot \|x'(n)\|^2 + \epsilon}, \tag{4}$$

where b is a compression factor that adjusts the degree of compression of the residual error. ϵ is a small positive constant that overcomes the possible numerical instability when updating the weights. In our case, it is equal to 2.2204×10^{-16}. Using the steepest descent method, the gradient of the loss function with respect to the filter coefficients $w(n)$ is computed as:

$$\nabla J(n) = \frac{-2b\tanh(b(n)e(n))[1 - \tanh^2(b(n)e(n))] \cdot x'(n)}{2 \cdot \|x'(n)\|^2 + \epsilon}, \tag{5}$$

where $x'(n)$ represents the filtered reference signal. The update rule for the THFxNLMS algorithm is therefore given by:

$$w(n+1) = w(n) - \mu\psi(e(n))x'(n), \tag{6}$$

where $\psi(e(n)) = \frac{\tanh(b(n)e(n))[1-\tanh^2(b(n)e(n))]}{2\cdot\|x'(n)\|^2+\epsilon}$ acts as a robust transformation function. The error signal is expressed as:

$$e(n) = d(n) - s(n) * w^T(n)x(n), \tag{7}$$

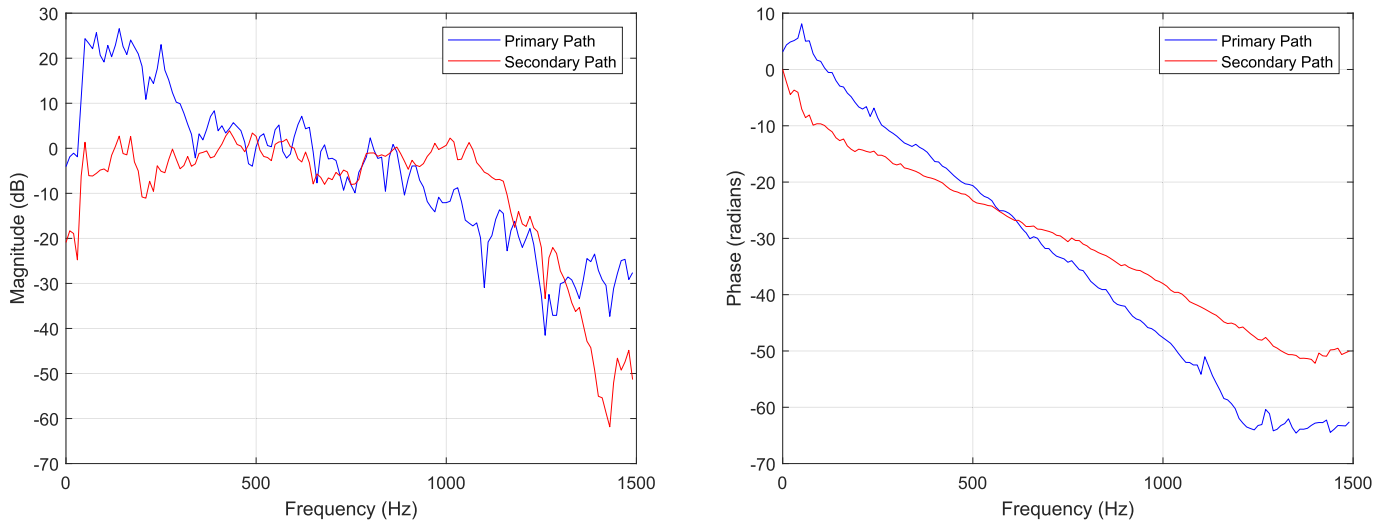
where $d(n)$ and $s(n)$ denote the primary disturbance and the impulse response of the secondary path, respectively, and $*$ represents the linear convolution. The stability of the THFxNLMS algorithm depends on the choice of the step size μ. Considering the weight update rule presented in Eq. 6. a sufficient condition to ensure stability is:

$$0 < \mu < \frac{1}{\lambda_{\max} D_s},$$

where:

- D_s is the group delay of the secondary path, which can be estimated from its impulse response $s(n)$;
- $\lambda_{\max}$ is the maximum eigenvalue of the autocorrelation matrix of the filtered reference signal $x'(n)$:

$$R_{x'} = E\left[x'(n)x'^T(n)\right], \quad \lambda_{\max} = \max \text{ eig}(R_{x'}).$$



(a) Magnitude response (b) Phase response

Fig. 2. Frequency responses of the primary and secondary paths identified using the NLMS algorithm.

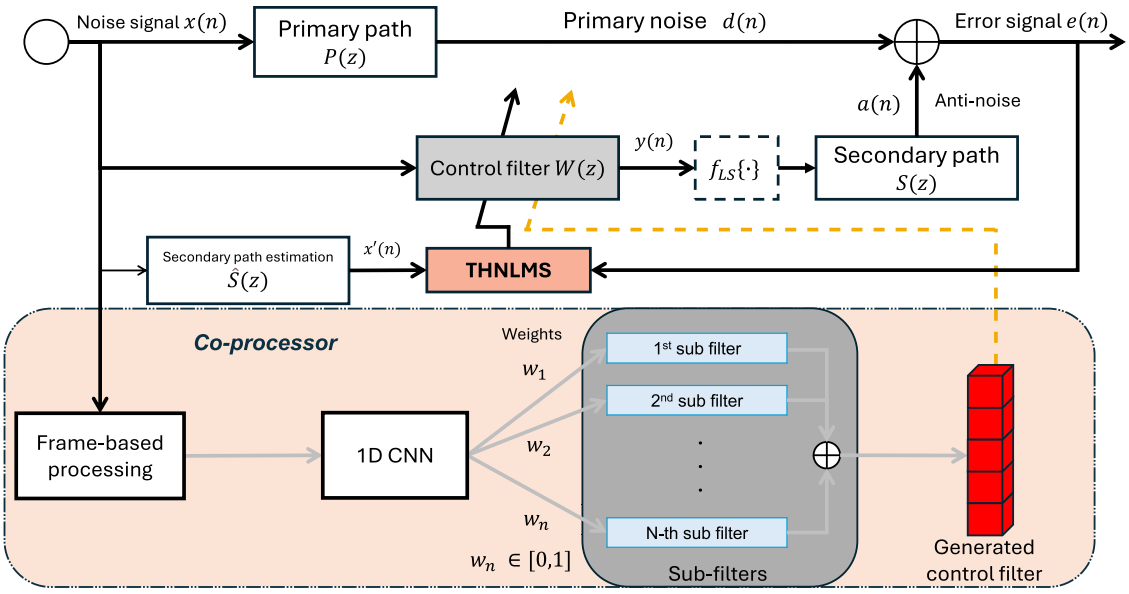


Fig. 3. Block diagram of the hybrid GFANC-THFxNLMS algorithm. f_{LS} represents the nonlinearity function associated with the loudspeaker. The noise signal is processed by the co-processor frame by frame, and for each frame, the 1D CNN provides the most suitable weight to design the control filter. This control filter is then fine-tuned by the THFxNLMS algorithm.

This inequality provides a verifiable upper bound for μ, ensuring that the adaptive update remains stable under the nonlinear transformation $\psi(e(n))$. It is important to note that $\lambda_{\max}$ depends on the reference signal statistics and D_s is determined by the secondary path. Therefore, in practical applications, μ is chosen conservatively based on the expected signal and path characteristics to ensure robust performance. The inclusion of the hyperbolic tangent function in the THFxNLMS algorithm enhances its robustness against nonlinearities, making it more suitable for ANC applications in complex acoustic environments where standard adaptive algorithms may struggle to achieve optimal performance. In addition, as indicated in (Li et al., 2023) the choice of the parameter b significantly influences the behavior of the robust transformation function. A higher b-value results in a steeper transformation curve, enhancing the algorithm's ability to suppress outliers. However, this also leads to a slower convergence rate. Conversely, a lower b-value produces a smoother transformation curve, accelerating the convergence speed but weakening the suppression of outliers. Additionally, a small b-value ensures that the step size remains stable during the adaptive process, preventing abrupt changes in the update rule.

In practical applications, selecting an appropriate b-value is crucial and should be based on the characteristics of the noise environment. To address the challenge of choosing b, an adaptive adjustment mechanism using a sigmoid function is introduced:

$$b(n) = \frac{\rho}{1 + \exp(-P_e(n))}, \tag{8}$$

$$P_e(n) = \lambda P_e(n-1) + (1-\lambda)e^2(n), \tag{9}$$

where ρ is a scaling factor within the range $[0, 1]$, $P_e(n)$ represents the energy estimation of the error signal, and λ is a forgetting factor, typically set within the range $0.9 < \lambda < 1$. This adaptive adjustment ensures that the transformation function dynamically adapts to

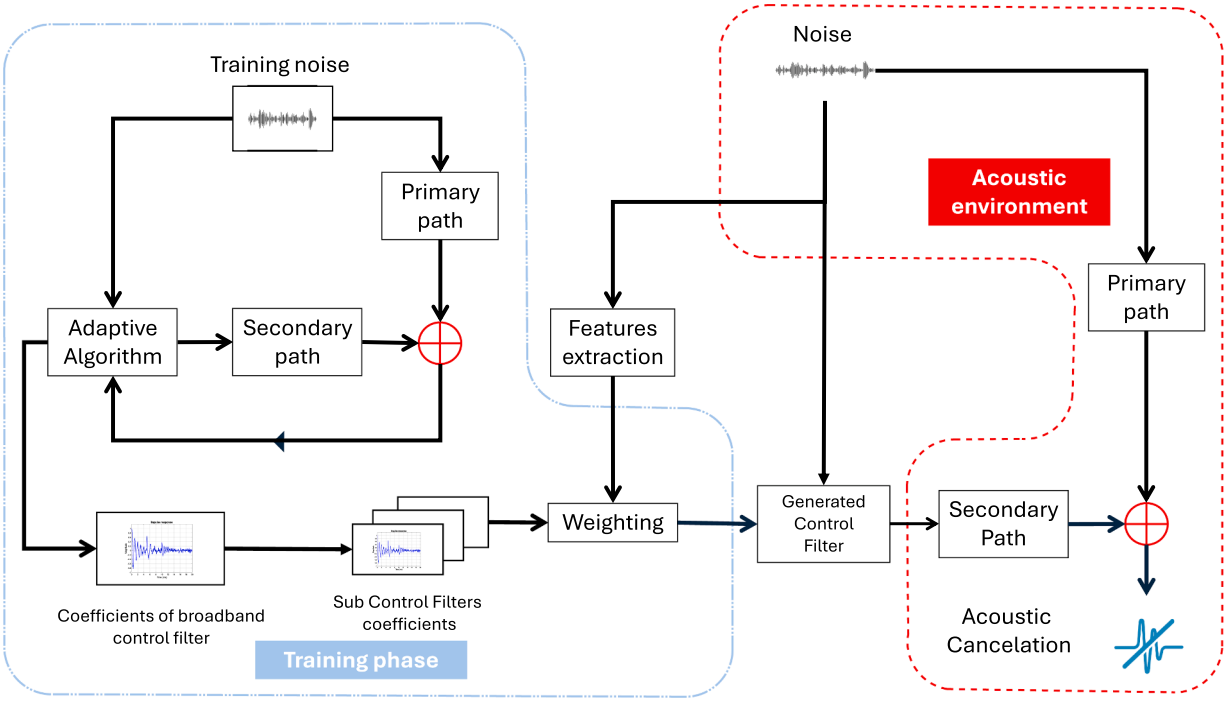


Fig. 4. Block diagram of the generative fixed-filter active noise control method, inspired by (Shi et al., 2020).

varying noise conditions, improving the robustness and stability of the THFxNLMS algorithm. A full derivation of THFxNLMS is proposed in Appendix B.

3.2. Convolutional neural network-Based generative fixed-Filter active noise control

3.2.1. GFANC Method

The Generative Fixed-Filter Active Noise Control (GFANC) method introduces a noise reduction approach by pre-training a set of broadband control filters. Based on this initial broadband control filter, a database of sub-control filters is created, allowing a weighting mechanism to dynamically generate an optimal control filter during operation (Luo et al., 2023). The method operates in two distinct stages: (1) offline filter pre-training and subdivision, and (2) real-time noise control, as illustrated in Fig. 4.

During the first phase, the target ANC system is used to cancel a broadband primary noise $d(n)$ containing the relevant frequency components. The optimal control filter, determined using the THFxNLMS algorithm, serves as the pre-trained broadband control filter. Once obtained, this filter is decomposed into sub-control filters using a practical filter decomposition technique based on the theory of perfect reconstruction filters (Dam et al., 2004). The decomposition of the broadband control filter relies on a perfect-reconstruction filter bank. Given a pre-trained control filter c with L coefficients, its frequency-domain representation is first obtained using the discrete Fourier transform (DFT):

$$G = F_L\, c$$

Where F_L is the DFT matrix. Exploiting the conjugate symmetry property of real signals in the frequency domain, the control filter is then decomposed into C sub-filters G_m, each covering a specific frequency range:

$$G = \sum_{m=1}^{C} G_m$$

Each sub-filter is defined such that its frequency components are extracted according to a perfect-reconstruction criterion, ensuring no loss of information. The time-domain representation of each sub-filter is then obtained through the inverse DFT:

$$c_m = F_L^{-1} G_m$$

Finally, the control signal can be reconstructed by summing the contributions of all sub-filters:

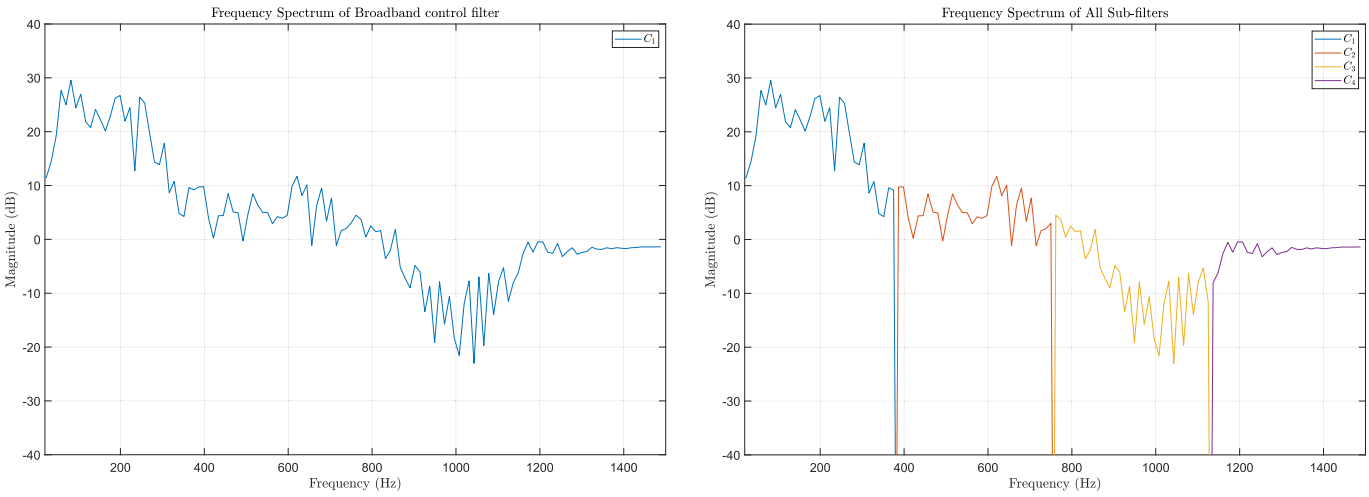
$$y(n) = \sum_{m=1}^{C} x^T(n) c_m$$

This decomposition method ensures that the original broadband filter is perfectly reconstructed in the time domain while allowing the generation of multiple fixed sub-filters. This process enhances flexibility in adaptive control strategies while preserving the optimal characteristics of the original filter (Luo et al., 2023). Resulting broadband filter and subfilters are presented in Fig. 5. They were computed with a sampling rate of 3 kHz and a control filter length set to 256 taps.

The 1D CNN does not model the noise directly, but learns the relationship between the spectral content of the reference signal and the optimal gain vectors for the sub-control filters. Based on these features, it generates continuous weights between 0 and 1 to combine the sub-filters and form the most suitable control filter for the input signal. This frequency-to-gain mapping clarifies that the CNN predicts optimal gains rather than the noise itself, justifying the use of synthetic training data while enabling generalization to real-world noises.

3.2.2. 1D Convolutional neural network architecture and training

Motivated by (Shi et al., 2020), the proposed 1D Convolutional Neural Network (CNN) is designed to process time-domain signals of 1-second duration, extracting relevant features to adaptively design the control filter. The input consists of a one-dimensional waveform, which is passed through a series of convolutional and fully connected layers to generate the optimal filter weights. The first convolutional layer applies 10 filters with a kernel size of 5, capturing local temporal dependencies in the signal. This is followed by a second convolutional layer with 20 filters and a larger kernel size of 32, allowing for deeper and more abstract feature extraction. A Max-Pooling operation with a kernel



(a) (b)

Fig. 5. Frequency responses of the broadband control filter (a) and the sub control filters (b).

Table 1
Architecture of the 1D CNN Model.

Layer	Type	Parameters	Output Shape
Conv1	Conv1D	In: 1, Out: 10, Kernel: 5, Stride: 1	(10, 15996)
Conv2	Conv1D	In: 10, Out: 20, Kernel: 32, Stride: 1	(20, 15965)
MaxPool	MaxPool1D	Kernel: 1024, Stride: 1024	(20, 15)
Flatten	Flatten	-	(40)
BatchNorm	BatchNorm1D	Num features: 40	(40)
Linear1	Fully Connected	In: 40, Out: 4	(4)
Activation	Tanh	-	(4)
Linear2	Fully Connected	In: 4, Out: 4	(4)

and stride size of 1024 is then applied to significantly reduce the temporal resolution while emphasizing the most informative components. The resulting feature map is flattened into a 40-dimensional vector and normalized using batch normalization to stabilize and accelerate training. The network then processes the extracted features through two fully connected layers: the first maps the features to a 4-dimensional space, followed by a Tanh activation function introducing non-linearity, and the second refines the output into the final set of 4 continuous values. These values correspond to the adaptive weights used to update the sub-control filters in real time. By leveraging hierarchical feature extraction and dimensionality reduction, this architecture enables the CNN to generate an effective and responsive noise control strategy, tailored to the spectral characteristics of the input signal. The detailed architecture is presented in Table 1.

The training process of the 1D CNN model is designed for a regression task, where the goal is to predict adaptive weights associated with input signals. The dataset consists of 80,000 training samples, each representing a 1-second temporal signal filtered within a randomly selected frequency range between 20 Hz and 1480 Hz. The corresponding target values are 4 continuous-valued coefficients ranging between 0 and 1, one for each subfilter. The validation set comprises 2000 samples following the same preprocessing pipeline. The model is trained using a Mean Squared Error (MSE) loss function, given the regression nature of the task. The optimizer employed is Adam, initialized with a learning rate of 0.0002 and a weight decay of 1e-4, ensuring both efficient convergence and regularization. A learning rate scheduler is also applied, reducing the learning rate by a factor of 0.5 every 5 epochs to facilitate stable optimization. The model is trained for 30 epochs using a batch size of 500, with training and validation losses monitored at each epoch. During training, the network undergoes weight initialization using Xavier initialization (Glorot & Bengio, 2010) for convolutional layers. The training loop consists of forward propagation, loss computation, backpropagation, and parameter updates, while validation is performed without gradient computation to assess generalization. In Fig. 6 the loss curves are presented.

Table 2
Computational complexity of the proposed 1D CNN model.

Component	Parameters	MACs per frame
1D CNN	6.74 K	19.33 M

To assess the real-time feasibility of the proposed framework, we analyze the computational complexity of the 1D CNN used for filter generation. The model contains a total of 6.74K trainable parameters, making it lightweight compared to typical deep ANC architectures. Despite its compact size, the convolutional layers account for most of the computational load. Using the `ptflops` library, the total number of Multiply-Accumulate operations (MACs) required for a single forward pass of a 1-second frame sampled at 3 kHz is measured as 19.33 million MACs. Table 2 summarizes the parameter count and MAC cost of the network.

Beyond raw computational complexity, memory access and buffer management are also critical factors for real-time deployment on hybrid embedded platforms such as the ADSP-SC598. Due to its small parameter footprint, the entire CNN model can be stored in on-chip memory, minimizing external memory access and associated latency. Moreover, the frame-based processing strategy allows the input buffer to be reused across successive inferences, while the generated filter coefficients are transferred only once to the real-time controller. This design limits memory bandwidth requirements and avoids frequent buffer switching, which is particularly important in heterogeneous CPU-DSP systems.

These results confirm that the proposed CNN architecture is suitable for embedded deployment.

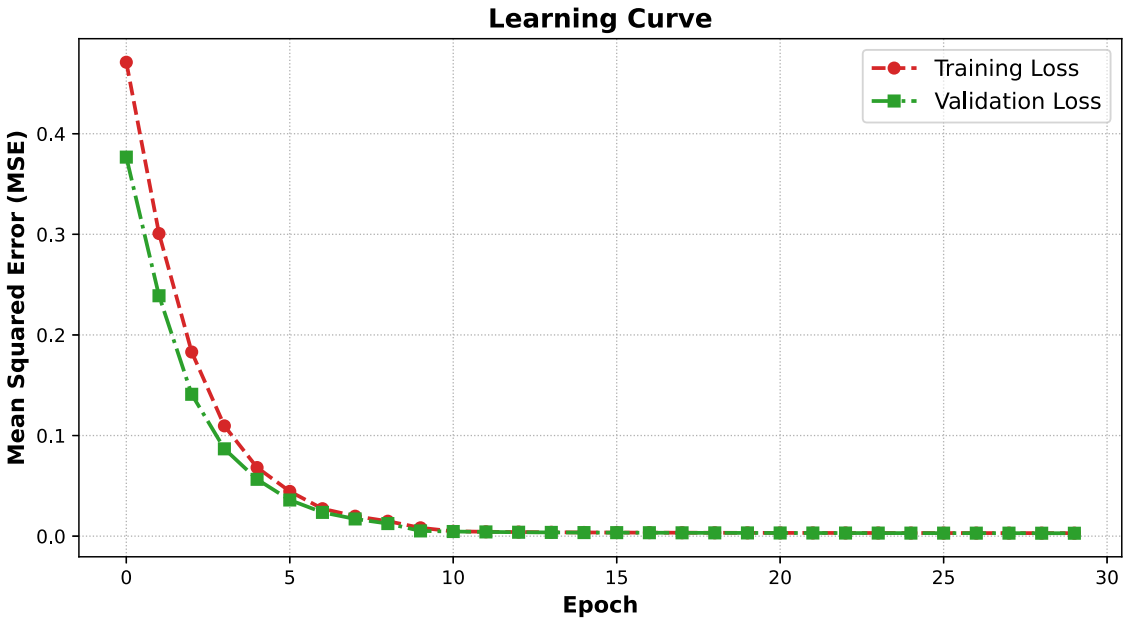


Fig. 6. Loss function curve of the 1D CNN training.

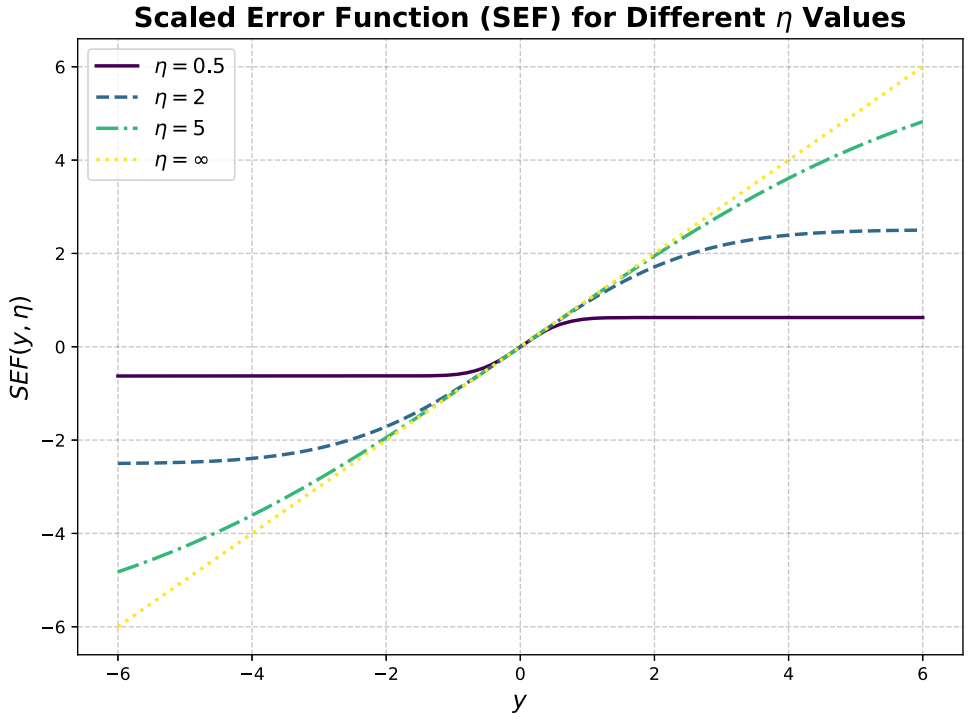


Fig. 7. Scaled error function (SEF) for different values of η.

3.2.3. Data construction

Concerning the data used to train the neural network, the dataset consisted, as previously mentioned, of 80,000 samples filtered with noise in a random frequency range for training and 2000 samples each for validation and testing. All noise signals were generated with a sampling rate of 3 kHz and a duration of 1 second. Specifically, the training samples consisted of narrow-band and broad-band noise signals with passbands randomly selected across the frequency range of interest. In the proposed approach, the labels used for supervised training correspond to adaptive gain vectors $\mathbf{w}_n^{(i)} \in \mathbb{R}^C$, which are estimated dynamically at each time step n for each sample i in a batch of size N, where C is the number of pre-trained control filters. In our case $N = 1000$ and $C = 4$.

The adaptation process follows a normalized least mean squares (NLMS) algorithm. Each control filter is fixed and stored in the matrix $\mathbf{H} \in \mathbb{R}^{C\times L}$, where L is the filter length. The input signal $x_n^{(i)}$ at time n is inserted into a rolling buffer $\mathbf{X}_d \in \mathbb{R}^{N\times L}$ that retains the past L samples for each signal in the batch. The output of the control filters is then computed as:

$$\mathbf{Y}_n = \mathbf{X}_d \cdot \mathbf{H}^\top \tag{10}$$

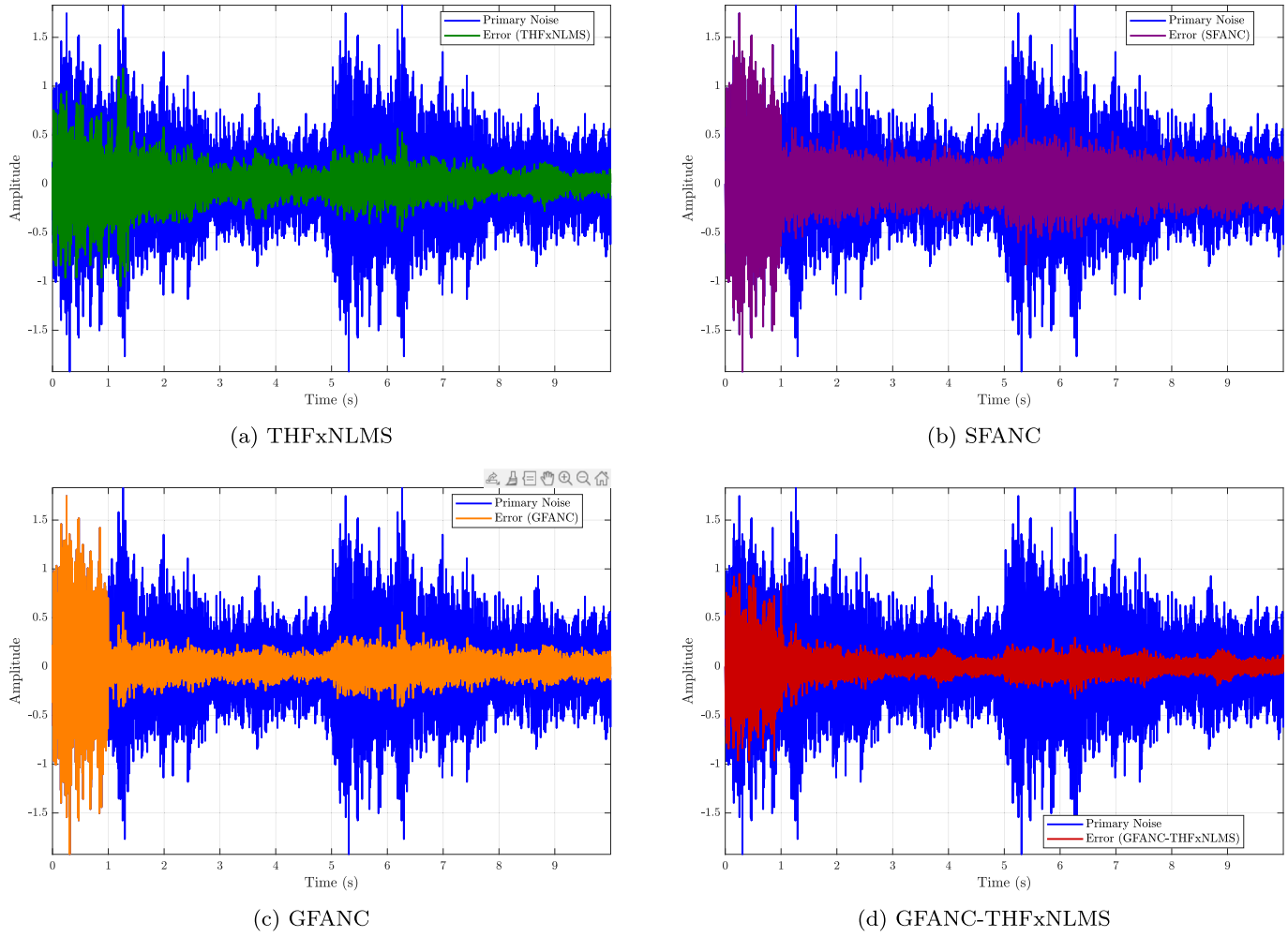


Fig. 8. Graphics of the primary noise and error signals obtained using different ANC algorithms. The control is applied to compressor noise with $\eta = 5$.

where each row $\mathbf{y}_n^{(i)} \in \mathbb{R}^C$ of $\mathbf{Y}_n$ corresponds to the filtered signal for sample i. The anti-noise signal is formed by a weighted sum of the control filter outputs:

$$\hat{d}_n^{(i)} = \mathbf{w}_n^{(i)} \cdot \mathbf{y}_n^{(i)} = \sum_{c=1}^{C} w_n^{(i,c)} \cdot y_n^{(i,c)} \tag{11}$$

and the residual error is computed as:

$$e_n^{(i)} = d_n^{(i)} - \hat{d}_n^{(i)} \tag{12}$$

where $d_n^{(i)}$ is the target disturbance. The gain vector $\mathbf{w}_n^{(i)}$ is updated using the NLMS rule:

$$\mathbf{w}_{n+1}^{(i)} = \mathbf{w}_n^{(i)} + \mu \cdot \frac{e_n^{(i)}}{\|\mathbf{y}_n^{(i)}\|^2} \cdot \mathbf{y}_n^{(i)} \tag{13}$$

with μ denoting the adaptation step size, and $\|\mathbf{y}_n^{(i)}\|^2 = \sum_{c=1}^{C} \left(y_n^{(i,c)}\right)^2$ representing the instantaneous output energy of the control filters. In our case $\mu = 0.001$.

At each iteration, the updated gain vector $\mathbf{w}_n^{(i)}$ represents the optimal linear combination (in the mean-square sense) of the control filter outputs. These vectors are recorded as training labels and paired with the corresponding system state $\mathbf{z}_n^{(i)}$, forming a dataset $\left\{\left(\mathbf{z}_n^{(i)}, \mathbf{w}_n^{(i)}\right)\right\}_{n,i}$ that is later used to train a predictive model capable of estimating $\mathbf{w}_n^{(i)}$ directly from the observed system state.

The objective of the 1D CNN is not to directly model real-world noise sources, but to learn how to optimally combine a set of pre-designed subfilters according to the spectral content of the reference signal. The network is therefore trained using synthetic band-pass noises with randomly generated passbands, exposing it to a wide range of controlled spectral profiles. For each training sample, the NLMS algorithm provides the optimal gain vector, which serves as the supervision signal and encodes the relationship between spectral content and subfilter weighting. As a result, the CNN learns a frequency-to-gain mapping that generalizes beyond the synthetic training data. Although real-world noises such as washing machines, engines, compressors, or airflow are more complex and non-stationary, their spectral characteristics remain within the range covered during training, allowing the CNN to generate appropriate subfilter weights for unseen real-world disturbances.

3.3. GFANC-THFxNLMS approach

The system processes incoming noise by sampling it at 3 kHz, capturing one second of audio per frame. As illustrated in Fig. 3, a dedicated coprocessor analyzes each frame using the GFANC framework. In this configuration, a one-dimensional convolutional neural network (1D CNN) is employed to generate a set of weights. These CNN-predicted weights are used to construct an initial optimal control filter that is tailored to the spectral characteristics of the current noise frame. Once the control filter is generated, it is transferred to the real-time controller, where the THFxNLMS algorithm takes over. Importantly, is to fine-tune the coefficients of the generated control filter during continuous operation. THFxNLMS is specifically designed to handle system nonlinearities through its nonlinear error formulation and adaptive

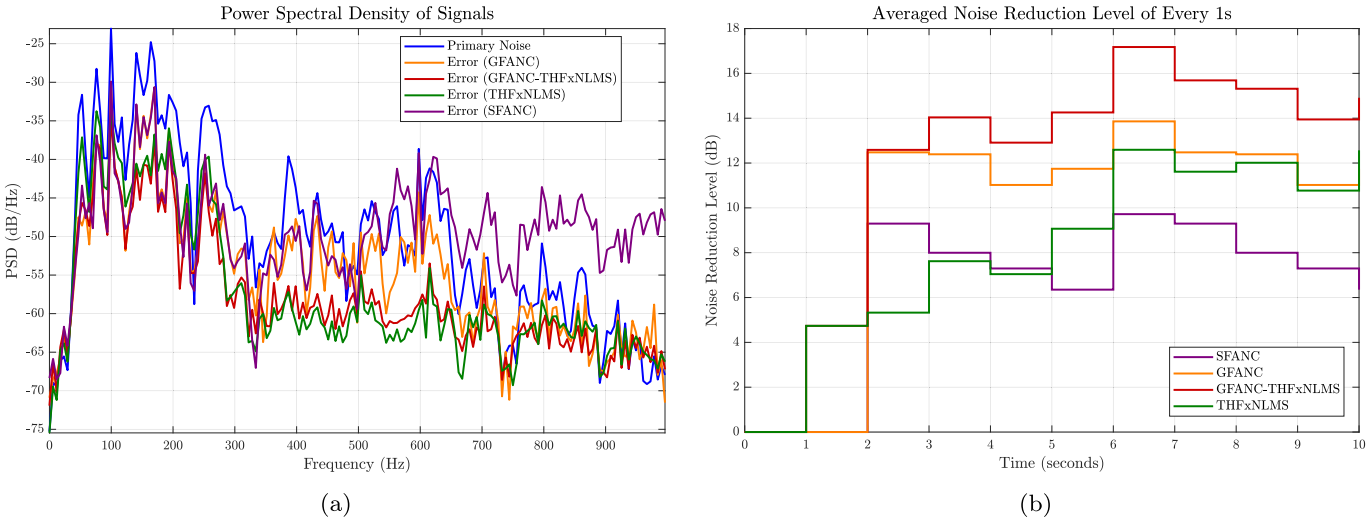


Fig. 9. Comparison of (a) PSD and (b) noise reduction per second obtained with THFxNLMS (green), GFANC (orange), SFANC (violet) and GFANC-THFxNLMS (red) with a compressor noise as a disturbance and $\eta = 5$.

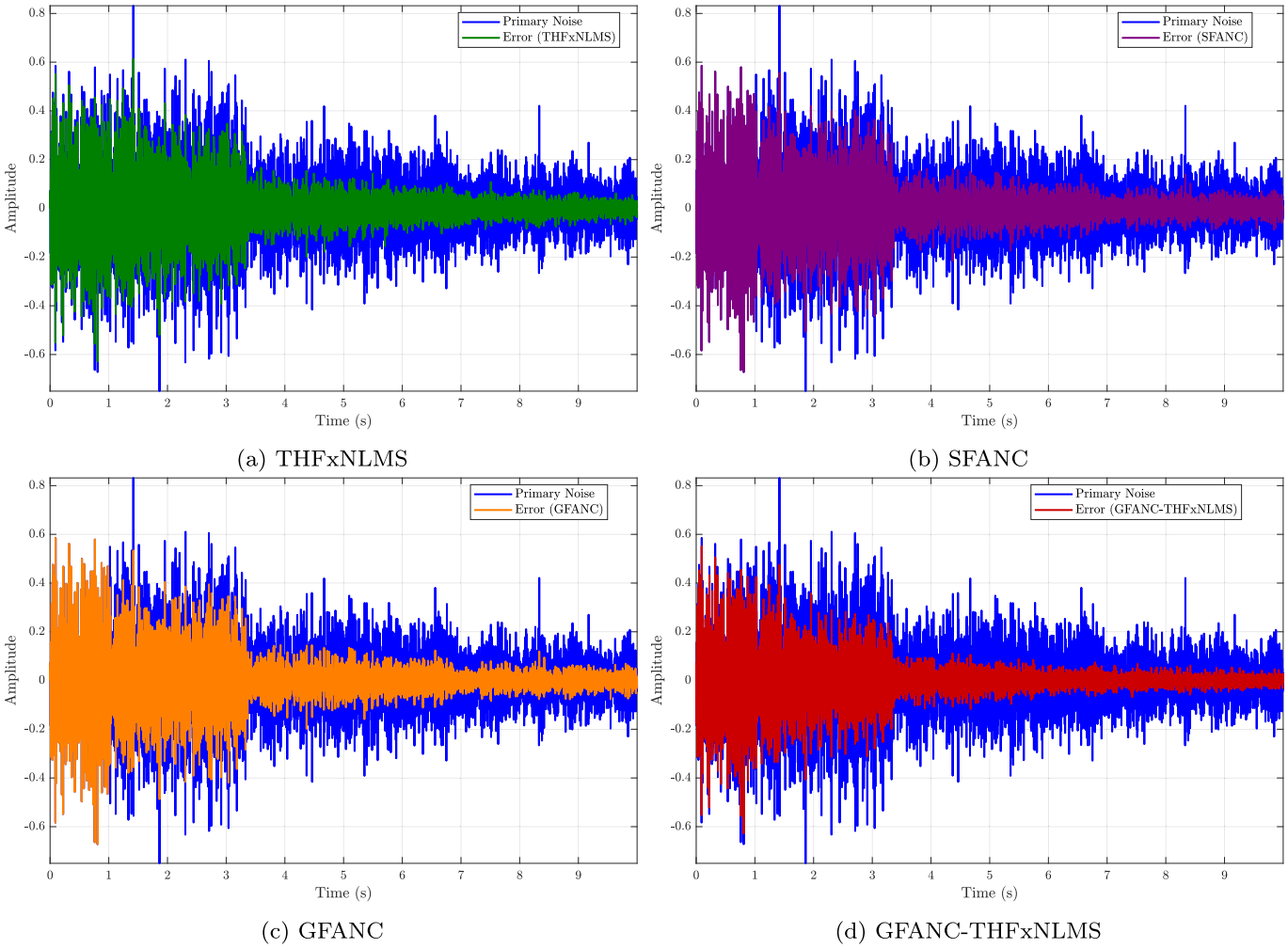


Fig. 10. Graphs of the primary noise and error signals obtained using different ANC algorithms. The control is applied to hood fan noise with $\eta = 5$, while the SNR varies over time.

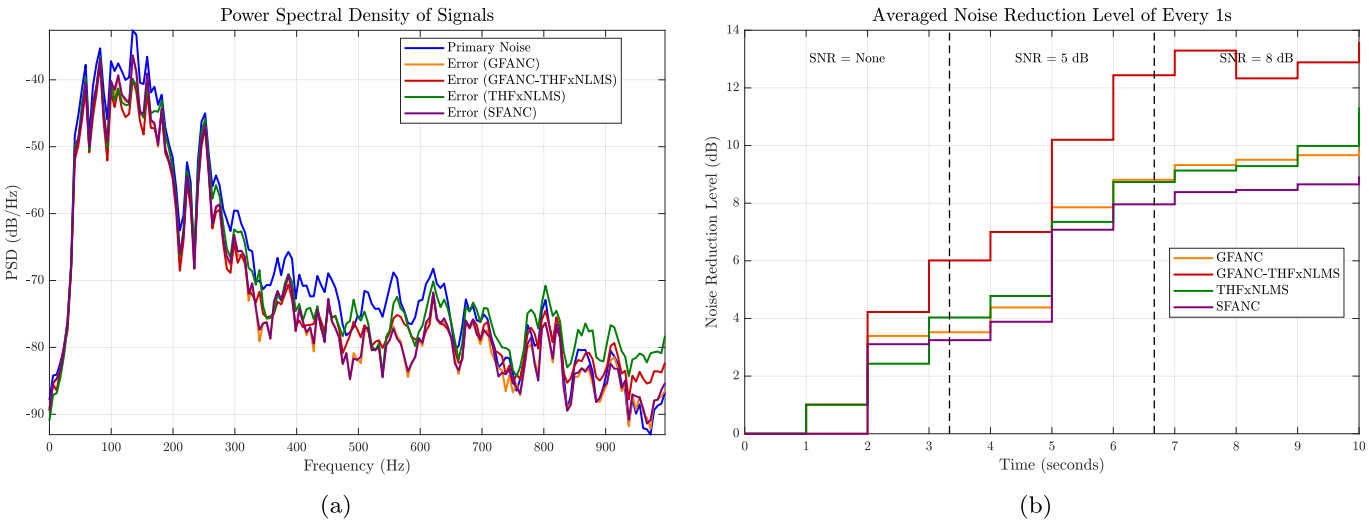


Fig. 11. Comparison of (a) PSD and (b) noise reduction per second obtained with THFxNLMS (green), GFANC (orange), SFANC (violet) and GFANC-THFxNLMS (red) with a hood fan noise as a disturbance and $\eta = 5$, while the SNR varies over time.

Table 3
Step-by-step description of the GFANC-THFxNLMS algorithm with corresponding equations.

Step Description	Corresponding Equation
1D CNN predicts weights	$\hat{\mathbf{w}}_{\text{CNN}} = \text{CNN}(x(n))$
Generated control filter	$\mathbf{w} = \hat{\mathbf{w}}_{\text{CNN}} \times \mathbf{H}$
Filtered reference signal	$\mathbf{x}'(n) = \hat{s}(n) * x(n)$
Compute controller output	$y(n) = \mathbf{w}^T(n)\mathbf{x}(n)$
Compute error signal	$e(n) = d(n) - s(n) * f_{SEF}(y(n))$
Update error power estimate	$P_e(n) = \lambda P_e(n-1) + (1-\lambda)e^2(n)$
Compute gradient scaling factor	$b(n) = \frac{\rho}{1+e^{-P_e(n)}}$
Compute weighted loss	$\text{loss}(n) = \frac{\tanh^2\left(b(n)e(n)\right)}{2 \cdot \|\mathbf{x}'(n)\|^2 + \varepsilon}$
Update filter weights	$\mathbf{w}(n+1) = \mathbf{w}(n) - \mu\psi(e(n))\mathbf{x}'(n)$
	where $\psi(e) = \frac{\tanh(b(n)e(n))[1-\tanh^2(b(n)e(n))]}{2\|\mathbf{x}'(n)\|^2 + \varepsilon}$

gradient scaling, allowing the controller to compensate for secondary-path distortions and nonlinear loudspeaker-acoustic effects that cannot be captured at the frame level. This separation of roles is central to the proposed architecture: the CNN operates at the frame level to provide a robust and noise dependent initialization of the control filter, while THFxNLMS performs sample level adaptation to explicitly address nonlinearities and time-varying dynamics in the ANC system. By combining the generative fixed filter approach with the THFxNLMS algorithm, the system ensures dynamic adaptation, progressively improving the noise reduction performance in real time. Unlike the algorithm proposed in (Luo et al., 2022), once the initial control filter is generated, THFxNLMS continues to adapt its coefficients throughout the remainder of the signal without requiring frame-wise CNN updates. Table 3 summarizes the step-by-step procedure of the proposed algorithm, where $\mathbf{H}$ denotes the subfilters dataset and $\hat{\mathbf{w}}_{\text{CNN}}$ represents the weights predicted by the 1D CNN.

4. Results

This section presents simulations using a single-channel ANC system. The setup includes a 3 kHz sampling rate and a control filter with 256 taps. The impulse responses previously introduced define the primary and secondary paths. Additionally, 4 pre-trained sub-control filters are employed. The THFxNLMS algorithm is adopted to derive the optimal control filters with a stepsize of 0.1. Loudspeaker saturation is a primary source of nonlinearity in ANC systems (Ghasemi et al., 2016). Several studies have examined this phenomenon (Cha et al., 2023; Zhang & Wang, 2021), often modeling it using the scaled error function (SEF) (Tobias & Seara, 2006):

$$f_{\text{SEF}}(y) = \int_0^y e^{-\frac{2z}{\eta^2}}\, dz, \tag{14}$$

where y is the loudspeaker input, and η determines the nonlinearity severity. This function represents a commonly observed saturation effect, such as sound level limitation due to loudspeaker size constraints. As $\eta \to \infty$, the SEF behaves linearly, while it acts as a hard limiter when $\eta \to 0$. To assess the robustness of the proposed approach against nonlinear distortions, three loudspeaker functions are considered during training:

- $\eta = 0.5$ (severe nonlinearity),
- $\eta = 5$ (moderate nonlinearity),
- $\eta = \infty$ (linear case).

Fig. 7 highlights the trade-off introduced by the nonlinearity parameter η. Large values of η correspond to quasi-linear loudspeaker behavior, where noise cancellation is mainly limited by the adaptive filter performance. For moderate values (e.g., $\eta = 5$), saturation effects are present but remain smooth, representing a realistic operating range in which the proposed method achieves a favorable balance between robustness and control accuracy. In contrast, very small values of η (e.g., $\eta = 0.5$) introduce severe saturation, where nonlinearity becomes a dominant factor and significantly challenges conventional ANC algorithms. Fig. 7 visualizes the SEF for these different η values. The evaluation phase includes testing with both trained and untrained loudspeaker functions. It is important to note that the pre-trained sub-control filters were obtained using training noise signals that were passed through nonlinear loudspeaker functions. This ensures that each sub-filter already incorporates robustness to saturation-type nonlinearities. The 1D CNN then efficiently combines these nonlinear-aware sub-filters, while the online THFxNLMS stage further fine-tunes the generated control filter coefficients.

To evaluate the efficiency of the algorithm, non-stationary noise was generated from real-life recordings, and control was computed under different degrees of nonlinearity. Additionally, a study was conducted to assess control performance under time-varying primary paths, examining the algorithm's robustness with moderate nonlinearities ($\eta = 5$). The results obtained with GFANC-THFxNLMS are compared to those

Table 4
Average Noise Reduction (in dB) for GFANC-THFxNLMS, GFANC, THFxNLMS, and SFANC with respect to different noises and nonlinear distortions.

Noise type	Engine				Compressor				Washing Machine				Air Flow			
η	∞	5	2	0.5	∞	5	2	0.5	∞	5	2	0.5	∞	5	2	0.5
THFxNLMS	8.85	8.82	8.68	5.96	7.76	7.74	7.66	6.03	9.60	8.86	8.12	5.16	6.74	6.74	6.73	6.52
SFANC	7.22	7.10	7.02	5.22	5.06	5.90	5.20	4.78	7.66	7.98	5.30	4.96	6.74	6.74	6.77	7.02
GFANC	7.41	7.40	7.32	5.33	6.14	6.14	6.11	5.11	7.95	7.60	7.50	5.35	8.01	8.01	8.00	7.79
GFANC-THFxNLMS	**11.44**	**11.42**	**11.30**	**7.07**	**10.27**	**10.25**	**10.16**	**7.60**	**10.64**	**10.18**	**9.12**	**6.41**	**9.74**	**9.73**	**9.73**	**9.49**

Table 5
Complexity analysis of the THFxNLMS algorithm.

Operation	Dimensions	$*$ Multiplications	$+ / -$ Additions/Subtractions
$x_f(n) = \hat{s}(n) * x(n)$	$\mathbb{R}^{L\times 1}$	L	1
$y(n) = \mathbf{w}^T(n)\mathbf{x}(n)$	$\mathbb{R}^{L\times 1}$	L	$L-1$
$e(n) = d(n) - s(n) * f_{SEF}(y(n))$	$\mathbb{R}$	1	1
$P_e(n) = \lambda P_e(n-1) + (1-\lambda)e^2(n)$	Scalar	2	1
$b(n) = \frac{\rho}{1+e^{-P_e(n)}}$	Scalar	2	1
$\text{loss}(n) = \frac{\tanh^2\big(b(n)e(n)\big)}{2\cdot\|\mathbf{x}'(n)\|^2+\varepsilon}$	Scalar	6	3
$\mathbf{w}(n+1) = \mathbf{w}(n) - \mu\psi(e(n))\mathbf{x}'(n), \quad \psi(e) = \frac{\tanh(b(n)e(n))\big[1-\tanh^2(b(n)e(n))\big]}{2\|\mathbf{x}'(n)\|^2+\varepsilon}$	$\mathbb{R}^{L\times 1}$	$4L+2$	$3L+1$
Total		$8L+11$	$4L+7$

of the THFxNLMS, Selective Fixed Filter Active Noise Control (SFANC) presented in (Shi et al., 2022) and GFANC algorithms.

4.1. ANC performance

To evaluate the algorithm's efficiency in noise control within the noise box, it was first tested using non-stationary noise sources, including sounds from a washing machine, engine, compressor, and airflow. Fig. 8 shows the results for washing machine noise. Fig. 8 presents signals in time domain and Fig. 9 illustrates the Power Spectral Density (PSD) of the primary noise and the controlled signals for all algorithms, as well as the average noise reduction achieved per second by each algorithm.

Fig. 8 illustrates the performance of various algorithms, highlighting that GFANC outperforms THFxNLMS for this noise type. This advantage significantly influences the effectiveness of GFANC-THFxNLMS. Notably, GFANC-THFxNLMS achieves lower error amplitudes due to its adaptive behavior. Additionally, GFANC introduces control after a delay of 1 second, allowing the CNN to analyze primary noise characteristics and determine the most appropriate controller. In contrast, GFANC-THFxNLMS applies control immediately, leveraging its adaptive properties.

Fig. 9 displays the Power Spectral Density (PSD) of each controlled signal up to 1000 Hz. The GFANC-THFxNLMS-controlled signal (red curve) exhibits the lowest levels. To evaluate its performance, it can be compared to the PSD of the GFANC controlled signal (yellow curve). In the 50–300 Hz range, the red curve remains either below or at the same level as the yellow one. Overall, GFANC-THFxNLMS demonstrates superior performance across most frequencies up to 1000 Hz, particularly in handling resonances and anti-resonances.

Fig. 9 illustrates the noise reduction achieved over time, measured at each second. The results show that GFANC and GFANC-THFxNLMS reach optimal noise reduction after 2 seconds. After 10 seconds of computation, THFxNLMS achieves a noise reduction similar to that of GFANC. This figure highlights two key aspects: first, the faster convergence time provided by GFANC, and second, the potential for improved noise reduction when effectively combining the characteristics of both algorithms. SFANC present an interesting convergence speed but obtain a lower noise reduction than GFANC. Table 4 presents the results obtained with different algorithms, noise types, and non-linearity conditions.

Table 4 reports the average noise reduction achieved by the three algorithms in different noise types and levels of nonlinear distortion. In all cases, the hybrid approach GFANC-THFxNLMS consistently outperforms both GFANC, SFANC and THFxNLMS individually, demonstrating the benefit of combining both methods. It is worth noting that GFANC occasionally yields lower noise reduction than THFxNLMS, which can be attributed to the initial one-second delay required by GFANC before it starts applying control, resulting in temporarily reduced effectiveness at the beginning of the signal.

4.2. ANC robustness

To assess the algorithm's robustness, a 10-second hood fan signal was used, with the SNR of the primary path varying over time. The SNR remained constant during the first third of the signal, was set to 5 dB for the second third, and then increased to 8 dB in the final third. Fig. 10 presents the results in the time domain, while Figs. 11a and 11b display the PSD of the controlled signals for each algorithm and the noise reduction achieved per second, respectively.

Regarding time-domain signals, the control performance appears similar across the three algorithms, but the lowest error amplitude is observed with GFANC-THFxNLMS. This observation is further supported by Fig. 11, which shows that GFANC-THFxNLMS achieves the best attenuation. Similarly, Fig. 11b clearly demonstrates that GFANC-THFxNLMS provides the most effective noise reduction. In the frequency domain, it is evident that the control achieved with GFANC-THFxNLMS is superior at low frequencies, particularly between 100 and 350 Hz. Regarding noise reduction per second, both in time-invariant and time-varying cases, it can be observed that, over time, the control obtained with THFxNLMS converges to that of GFANC. With longer-duration signals, THFxNLMS could potentially surpass GFANC. This highlights the benefit of merging GFANC and THFxNLMS, leveraging the strengths of both algorithms in terms of convergence speed and adaptability. However, a limitation in time-varying systems is that if the SNR variation becomes more significant, the control performance may degrade. The average noise reduction achieved by the algorithms is as follows: 3.63 dB for the THFxNLMS algorithm, 3.35 dB for GFANC, 3.12 dB for SFANC and 4.83 dB for GFANC-THFxNLMS.

Table 6
Complexity analysis of the SFANC algorithm. The 1D CNN outputs a scalar index identifying the most appropriate control filter. C_{CNN} denotes the computational cost of the CNN, and L is the filter length.

Operation	Dimensions	$*$ Multiplications	$+ / -$ Additions/Subtractions
1D CNN (scalar index prediction)	Input: $\mathbb{R}^{1\times1\times f_s}$	C_{CNN}	C_{CNN}
Filter selection: $\mathbf{w} = \mathbf{W}_c[i]$	$i \in \mathbb{R}$ (scalar index)	0	0
$x_f(n) = \hat{s}(n) * x(n)$	$\mathbb{R}^{L\times1}$	L	1
$y(n) = \mathbf{w}^T(n)\mathbf{x}(n)$	$\mathbb{R}^{L\times1}$	L	$L-1$
$e(n) = d(n) - s(n) * f_{SEF}(y(n))$	$\mathbb{R}$	1	1
Total		$C_{\text{CNN}} + 2L + 2$	$C_{\text{CNN}} + L + 2$

Table 7
Complexity analysis of the GFANC algorithm. $\hat{\mathbf{w}}_{\text{CNN}}$ denotes the weights predicted by the 1D CNN, C_{CNN} is its computational cost, and $\mathbf{H}$ represents the subfilters dataset. C represents the number of subfilters, and L represents the length of each filter.

Operation	Dimensions	$*$ Multiplications	$+ / -$ Additions/Subtractions
1D CNN	Input: $\mathbb{R}^{1\times1\times f_s}$	C_{CNN}	C_{CNN}
$\mathbf{w} = \hat{\mathbf{w}}_{\text{CNN}} \times \mathbf{H}$	$\mathbb{R}^{1\times C} \times \mathbb{R}^{C\times L}$	$C \times L$	$C \times (L-1)$
$x_f(n) = \hat{s}(n) * x(n)$	$\mathbb{R}^{L\times1}$	L	1
$y(n) = \mathbf{w}^T(n)\mathbf{x}(n)$	$\mathbb{R}^{L\times1}$	L	$L-1$
$e(n) = d(n) - s(n) * f_{SEF}(y(n))$	$\mathbb{R}$	1	1
Total		$C_{\text{CNN}} + CL + 2L + 1$	$C_{\text{CNN}} + C(L-1) + L + 1$

Table 8
Complexity analysis of the GFANC-THFxNLMS algorithm. $\hat{\mathbf{w}}_{\text{CNN}}$ denotes the weights predicted by the 1D CNN, C_{CNN} is its computational cost, and $\mathbf{H}$ represents the subfilters dataset. C represents the number of subfilters, and L represents the length of each filter.

Operation	Dimensions	$*$ Multiplications	$+ / -$ Additions/Subtractions
1D CNN	Input: $\mathbb{R}^{1\times1\times f_s}$	C_{CNN}	C_{CNN}
$\mathbf{w} = \hat{\mathbf{w}}_{\text{CNN}} \times \mathbf{H}$	$\mathbb{R}^{1\times C} \times \mathbb{R}^{C\times L}$	$C \times L$	$C \times (L-1)$
$x_f(n) = \hat{s}(n) * x(n)$	$\mathbb{R}^{L\times1}$	L	1
$y(n) = \mathbf{w}^T(n)\mathbf{x}(n)$	$\mathbb{R}^{L\times1}$	L	$L-1$
$e(n) = d(n) - s(n) * f_{SEF}(y(n))$	$\mathbb{R}$	1	1
$P_e(n) = \lambda P_e(n-1) + (1-\lambda)e^2(n)$	Scalar	2	1
$b(n) = \frac{\rho}{1+e^{-P_e(n)}}$	Scalar	2	1
$\text{loss}(n) = \frac{\tanh^2\left(b(n)e(n)\right)}{2\cdot\|\mathbf{x}'(n)\|^2+\varepsilon}$	Scalar	6	3
$\mathbf{w}(n+1) = \mathbf{w}(n) - \mu\psi(e(n))\mathbf{x}'(n), \quad \psi(e) = \frac{\tanh(b(n)e(n))\left[1-\tanh^2(b(n)e(n))\right]}{2\|\mathbf{x}'(n)\|^2+\varepsilon}$	$\mathbb{R}^{L\times1}$	$4L + 2$	$3L + 1$
Total		$C_{\text{CNN}} + CL + 8L + 11$	$C_{\text{CNN}} + CL + 4L + 8$

4.3. Computational complexity

In this section, we present a comparative analysis of the computational complexity of THFxNLMS, SFANC, GFANC, and GFANC-THFxNLMS. The computational complexity of the four algorithms is detailed in Tables 5–8. The THFxNLMS algorithm exhibits a relatively moderate complexity, with $8L + 11$ multiplications and $4L + 7$ additions/subtractions per iteration. The SFANC algorithm also integrates a pre-trained 1D CNN, its output is a scalar index to select the most appropriate pre-trained control filter. The total complexity per sample of SFANC amounts to $C_{\text{CNN}} + 2L + 2$ multiplications and $C_{\text{CNN}} + L + 2$ additions/subtractions. The GFANC algorithm integrates a pre-trained 1D CNN to predict weights to generate the control filter. Although the network is not updated during operation, the inference step still introduces a non-negligible computational cost called C_{CNN}. Combined with the matrix multiplication $\hat{\mathbf{w}}_{\text{CNN}} \times \mathbf{H}$, the total complexity amounts to $C_{\text{CNN}} + CL + 2L + 1$ multiplications and $C_{\text{CNN}} + C(L-1) + L + 1$ additions/subtractions. The GFANC-THFxNLMS algorithm combines the CNN-based subfilter generation with the adaptive refinement of the THFxNLMS update. As a result, its complexity is higher, with $C_{\text{CNN}} + CL + 8L + 11$ multiplications and $C_{\text{CNN}} + CL + 4L + 8$ additions/subtractions. This hybrid structure increases the computational load but leads to improved noise reduction performance compared to each component used individually. Although the number of MACs for the CNN inference is relatively low (19.33 M per frame), memory transfers for the CNN weights ($\hat{\mathbf{w}}_{\text{CNN}}$) and buffer management for frame-based processing are critical in real hardware. Efficient on-chip storage and reuse of input buffers help minimize memory latency and bandwidth issues in embedded implementations.

5. Conclusion

This study applies active noise control in an encapsulated structure with an opening using the GFANC-THFxNLMS algorithm. The GFANC-THFxNLMS algorithm follows a deep learning-based approach designed to enhance active noise control. The core strategy involves creating a dataset of multiple pre-trained sub-control filters derived from a broadband control filter. A 1D CNN dynamically generates the most suitable controller for the target noise, while an adaptive algorithm fine-tunes the coefficients of the chosen filter in real time. The primary goal of this algorithm is twofold: to improve efficiency compared to traditional adaptive algorithms and to address the limitations of GFANC in handling nonlinearities, particularly those introduced by the loudspeaker. These nonlinearities, which become more pronounced at higher amplitudes, can degrade control performance, making a more flexible and adaptive approach necessary. The results demonstrate that in an environment where such distortions are present, GFANC-THFxNLMS outperforms both GFANC and THFxNLMS. Specifically, Table 4 shows that GFANC-THFxNLMS achieves an average noise reduction of up to 11.44 dB for Engine noise, representing an improvement of approximately 2.59 dB over THFxNLMS, 4.22 dB over SFANC and 4.03 dB over GFANC. For

other noise types, similar numerical improvements are observed, e.g., a gain of 3.13 dB over THFxNLMS for compressor noise and 2.69 dB for washing machine noise. This hybrid approach enhances adaptability and convergence, ensuring effective noise reduction even under non-ideal conditions. However, the algorithm has certain limitations: when variations in the primary path coincide with significant SNR changes, control effectiveness diminishes, highlighting areas for further improvement. Moreover, when nonlinearities become too strong, the achievable noise reduction may be significantly limited, regardless of the control strategy used.

Looking ahead, deep learning could evolve from a supporting tool into the primary driver of control, directly generating the anti-noise signal. This shift would be particularly valuable for managing nonlinearities related to the environment and electroacoustic phenomena associated with the materials used in noise control systems. However, several challenges remain for deploying such systems in real-time applications. One major limitation concerns the large number of parameters in deep architectures, which can conflict with the very low-latency requirements of active noise control. In addition, the adaptive nature of ANC must be preserved, meaning that the algorithm should be able to adjust dynamically to new and unseen acoustic situations that were not encountered during training. These aspects highlight the need for lightweight and adaptive architectures capable of maintaining stability and real-time performance while dealing with the inherent variability of acoustic environments.

CRediT authorship contribution statement

Alkahf Aboutiman: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization; **Khaled Said Ahmed Maamoun:** Conceptualization, Validation, Formal analysis, Resources, Data curation, Writing – review & editing; **Hamid Reza Karimi:** Conceptualization, Validation, Writing – original draft, Writing – review & editing, Supervision; **Francesco Ripamonti:** Conceptualization, Validation, Writing – original draft, Writing – review & editing, Supervision.

Data availability

Data will be made available on request.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The work was supported by the project IN-NOVA: Active reduction of noise transmitted into and from enclosures through encapsulated structures, which has received funding from the European Union's Horizon Europe programme under the Marie Sklodowska-Curie grant agreement no. 101073037.

Appendix A. Notation Table

Table A.1
Notation used throughout the manuscript.

Symbol	Description
L	Length of the control filter (number of taps)
C	Number of sub-filters in the GFANC or GFANC-THFxNLMS algorithms
G	Frequency-domain representation of the pre-trained control filter
$x(n)$	Reference signal vector at discrete time n
$x'(n)$	Filtered reference signal vector at discrete time n
$y(n)$	Output of the control filter at discrete time n
$e(n)$	Residual error signal at discrete time n
$d(n)$	Primary disturbance signal at discrete time n
$s(n)$	Impulse response of the secondary path at discrete time n
$\hat{s}(n)$	Estimate of the secondary path impulse response at discrete time n
$w(n)$	Coefficient vector of the control filter at discrete time n
$w(n+1)$	Updated control filter coefficient vector at discrete time n
$P(z)$	Transfer function of the primary path
$W(z)$	Transfer function of the control filter
$S(z)$	Transfer function of the secondary path
$f(x)$	Nonlinear transformation function (hyperbolic tangent) used in THFxNLMS
$f_{\text{SEF}}(y)$	Scaled Error Function modeling loudspeaker saturation: $f_{\text{SEF}}(y) = \int_0^y e^{-2z/\eta^2}\, dz$
η	Parameter determining the severity of loudspeaker nonlinearity in SEF
b	Compression factor for the robust transformation function
$b(n)$	Adaptive compression factor at discrete time n
$\psi(e(n))$	Robust transformation function used in the THFxNLMS weight update
$J(n)$	Loss function at discrete time n
$P_e(n)$	Error power estimate at discrete time n
λ	Forgetting factor used in the error power estimation
ρ	Scaling factor for the adaptive compression parameter $b(n)$
μ	Step size of the adaptive algorithm
ϵ	Small constant to ensure numerical stability during weight update
D_s	Group delay of the secondary path
$\lambda_{\max}$	Maximum eigenvalue of the autocorrelation matrix of $x'(n)$
C_{CNN}	Computational cost of the 1D CNN
$\mathbf{H} \in \mathbb{R}^{C\times L}$	Dataset of pre-trained sub-filters
$\mathbf{w}_n^{(i)} \in \mathbb{R}^C$	Gain vector for sample i at time n, used to weight sub-filters
$\mathbf{X}_d \in \mathbb{R}^{N\times L}$	Rolling buffer storing past L samples for each of N signals in a batch
$\mathbf{Y}_n \in \mathbb{R}^{N\times C}$	Outputs of all control filters at time n for all N samples
$\mathbf{y}_n^{(i)} \in \mathbb{R}^C$	Row vector of $\mathbf{Y}_n$ for sample i, filtered sub-filter outputs
$\hat{d}_n^{(i)}$	Anti-noise signal formed by weighted sum of sub-filter outputs for sample i at time n
$e_n^{(i)}$	Residual error for sample i at time n
N	Batch size for training the CNN model
$\mathbf{z}_n^{(i)}$	Observed system state for sample i at time n, used as input to predictive model

Appendix B. Full Derivation of the THFxNLMS Algorithm

The THFxNLMS algorithm defines the loss function as:

$$J(n) = \frac{f^2(b(n)e(n))}{2\|x'(n)\|^2 + \epsilon}, \tag{B.1}$$

where

$$f(x) = \tanh(x),$$

b is the compression factor, $x'(n)$ is the filtered reference signal, and ϵ is a small positive constant to prevent division by zero. The residual error is:

$$e(n) = d(n) - s(n) * w^T(n)x(n), \tag{B.2}$$

where $d(n)$ is the primary disturbance and $s(n)$ is the secondary path impulse response.

The gradient of the loss with respect to the filter coefficients $w(n)$ is computed using the chain rule. Let

$$g(e(n)) = f(b(n)e(n)) = \tanh(b(n)e(n)).$$

Then the derivative of $g^2(e(n))$ with respect to $e(n)$ is:

$$\frac{\partial g^2(e(n))}{\partial e(n)} = 2\, g(e(n))g'(e(n)) = 2\tanh(b(n)e(n)) \cdot b(n)\left[1 - \tanh^2(b(n)e(n))\right].$$

Applying the chain rule to the full loss function $J(n)$:

$$\begin{aligned}\nabla J(n) &= \frac{\partial J(n)}{\partial w(n)} = \frac{\partial J(n)}{\partial e(n)} \cdot \frac{\partial e(n)}{\partial w(n)} \\ &= \frac{2b(n)\tanh(b(n)e(n))\left[1 - \tanh^2(b(n)e(n))\right] \cdot (-x'(n))}{2\|x'(n)\|^2 + \epsilon}.\end{aligned} \tag{B.3}$$

We define the robust transformation function:

$$\psi(e(n)) = \frac{\tanh(b(n)e(n))\left[1 - \tanh^2(b(n)e(n))\right]}{\|x'(n)\|^2 + \epsilon}, \tag{B.4}$$

so that the weight update rule can be written as:

$$w(n+1) = w(n) - \mu\psi(e(n))x'(n). \tag{B.5}$$

To ensure stability of the adaptive update, the step size μ must satisfy:

$$0 < \mu < \frac{1}{\lambda_{\max} D_s}, \tag{B.6}$$

where

$$R_{x'} = E\left[x'(n)x'^T(n)\right], \quad \lambda_{\max} = \max \operatorname{eig}(R_{x'}), \tag{B.7}$$

is the maximum eigenvalue of the autocorrelation matrix of the filtered reference signal $x'(n)$, and D_s is the secondary path group delay. This condition guarantees that the nonlinear transformation $\psi(e(n))$ does not destabilize the weight update.

To improve robustness against nonlinearities and varying noise conditions, the compression factor b is adapted using:

$$b(n) = \frac{\rho}{1 + \exp(-P_e(n))}, \tag{B.8}$$

$$P_e(n) = \lambda P_e(n-1) + (1-\lambda)e^2(n), \tag{B.9}$$

where ρ is a scaling factor in $[0, 1]$ and λ is a forgetting factor. This ensures that $\psi(e(n))$ dynamically adjusts to the current error level, enhancing stability and robustness in varying noise conditions.

References

Aboutiman, A., Rachman, Z., Oberman, T., Aletta, F., Kang, J., Karimi, H. R., & Ripamonti, F. (2025a). Subjective perception analysis of active noise control algorithms in an encapsulated structure: An experimental study. *Applied Acoustics*, *239*, 110823. https://doi.org/10.1016/j.apacoust.2025.110823

Aboutiman, A., Shams, R., Karimi, H. R., Ripamonti, F., & Pawełczyk, M. (2025b). Active noise control in encapsulated structures with non-minimum phase characteristics using a kalman filter approach. *Journal of Sound and Vibration*, *615*, 119187. https://doi.org/10.1016/j.jsv.2025.119187

Analog Devices (2024). ADSP-SC598 Product Page. Accessed: 2024-12-29 https://www.analog.com/en/products/adsp-sc598.HTML.

Bambang, R. T. (2008). Adjoint EKF learning in recurrent neural networks for nonlinear active noise control. *Applied Soft Computing*, *8*(4), 1498–1504. https://doi.org/10.1016/j.asoc.2007.10.017

Bouchard, M., Paillard, B., & Le Dinh, C. T. (1999). Improved training of neural networks for the nonlinear active control of sound and vibration. *IEEE Transactions on Neural Networks*, *10*(2), 391–401. https://doi.org/10.1109/72.750568

Cha, Y.-J., Choi, W., Suh, G., Mahmoudkhani, S., & Büyüköztürk, O. (2018). Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types. *Computer-Aided Civil and Infrastructure Engineering*, *33*(9), 731–747. https://doi.org/10.1111/mice.12334

Cha, Y.-J., Mostafavi, A., & Benipal, S. S. (2023). Dnoisenet: Deep learning-based feedback active noise control in various noisy environments. *Engineering Applications of Artificial Intelligence*, *121*, 105971. https://doi.org/10.1016/j.engappai.2023.105971

Chang, C.-Y., & Luoh, F.-B. (2007). Enhancement of active noise control using neural-based filtered-x algorithm. *Journal of Sound and Vibration*, *305*(1–2), 348–356. https://doi.org/10.1016/j.jsv.2007.04.007

Dam, H. H., Nordholm, S., Cantoni, A., & de Haan, J. M. (2004). Iterative method for the design of DFT filter bank. *IEEE Transactions on Circuits and Systems II: Express Briefs*, *51*(11), 581–586. https://doi.org/10.1109/TCSII.2004.836041

Das, D. P., & Panda, G. (2004). Active mitigation of nonlinear noise processes using a novel filtered-s LMS algorithm. *IEEE Transactions on Speech and Audio Processing*, *12*(3), 313–322. https://doi.org/10.1109/TSA.2003.822741

Duan, Y., Yisheng, L. V., & Wang, F.-Y. (2016). Travel time prediction with LSTM neural network. In *2016 IEEE 19Th international conference on intelligent transportation systems (ITSC)* (pp. 1053–1058). IEEE. https://doi.org/10.1109/ITSC.2016.7795686

Elliott, S., Stothers, I., & Nelson, P. (1987). A multiple error LMS algorithm and its application to the active control of sound and vibration. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *35*(10), 1423–1434. https://doi.org/10.1109/TASSP.1987.1165044

Elzaghmouri, B. M., Jbara, Y. H. F., Elaiwat, S., Innab, N., Osman, A. A. F., Ataelfadiel, M. A. M., Zawaideh, F. H., Alawneh, M. F., Al-Khateeb, A., & Abu-Zanona, M. (2024). A novel hybrid architecture for superior iot threat detection through real iot environments. *Computers, Materials & Continua*, *81*(2).

Fu, R., Zhang, Z., & Li, L. (2016). Using LSTM and GRU neural network methods for traffic flow prediction. In *2016 31St youth academic annual conference of chinese association of automation (YAC)* (pp. 324–328). IEEE. https://doi.org/10.1109/YAC.2016.7804912

Fuller, C. R., Hansen, C. H., & Snyder, S. D. (1991). Active control of sound radiation from a vibrating rectangular panel by sound sources and vibration inputs: An experimental comparison. *Journal of Sound and Vibration*, *145*(2), 195–215. https://doi.org/10.1016/0022-460X(91)90587-A

Ghasemi, S., Kamil, R., & Marhaban, M. H. (2016). Nonlinear thf-fxlms algorithm for active noise control with loudspeaker nonlinearity. *Asian Journal of Control*, *18*(2), 502–513. https://doi.org/10.1002/asjc.1140

Glorot, X., & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics* (pp. 249–256). JMLR Workshop and Conference Proceedings.

Goodwin, G. C., Silva, E. I., & Quevedo, D. E. (2010). Analysis and design of networked control systems using the additive noise model methodology. *Asian Journal of Control*, *12*(4), 443–459. https://doi.org/10.1002/asjc.201

Guo, X., Li, Y., Jiang, J., Dong, C., Du, S., & Tan, L. (2018). Sparse modeling of nonlinear secondary path for nonlinear active noise control. *IEEE Transactions on Instrumentation and Measurement*, *67*(3), 482–496. https://doi.org/10.1109/TIM.2017.2781992

Hartmann, W. M. (2004). Signals, sound, and sensation. Springer Science & Business Media.

Isaac, C. W., Wrona, S., Pawelczyk, M. et al. (2022). Modelling vibro–acoustic response of lightweight square aluminium panel influenced by sound source locations for active control. *Scientific Reports*, *12*(1), 10727. https://doi.org/10.1038/s41598-022-14951-y

Jeong, I., & Park, Y. (2023). Suboptimal controller design of global active noise control system for various acoustic environments. *Scientific Reports*, *13*(1), 5453. https://doi.org/10.1038/s41598-023-32261-9

Jun, S., Ya-li, Z., & Qi-zhi, Z. (2012). Arc-tangent transformation algorithm for active impulsive noise control. *Noise and Vibration Control*, *32*(2), 27. https://doi.org/10.3969/j.issn.1006-1355-2012.02.006

Krukowicz, T. (2010). Active noise control algorithm based on a neural network and nonlinear input-output system identification model. *Archives of Acoustics*, *35*(2), 191–202.

Kuo, S. M., & Morgan, D. R. (1996). Active noise control systems (vol. 4). Wiley, New York.

Kuo, S. M., & Wu, H.-T. (2005). Nonlinear adaptive bilinear filters for active noise control systems. *IEEE Transactions on Circuits and Systems I: Regular Papers*, *52*(3), 617–624. https://doi.org/10.1109/TCSI.2004.842429

Lam, B., Gan, W.-S., Shi, D., Nishimura, M., & Elliott, S. (2021). Ten questions concerning active noise control in the built environment. *Building and Environment*, *200*, 107928. https://doi.org/10.1016/j.buildenv.2021.107928

Lam, B., Shi, D., Gan, W.-S., Elliott, S. J., & Nishimura, M. (2020). Active control of broadband sound through the open aperture of a full-sized domestic window. *Scientific Reports*, *10*(1), 10021. https://doi.org/10.1038/s41598-020-66563-z

Lashkari, K. (2006). A novel volterra-wiener model for equalization of loudspeaker distortions. In *2006 IEEE International conference on acoustics speech and signal processing proceedings* (pp. 1–5). IEEE (*vol. 5*). https://doi.org/10.1109/ICASSP.2006.1661226

Li, C., Jin, G., Liu, H., & Li, J. (2023). Active impulsive noise control algorithm based on adjustable hyperbolic tangent function. *Circuits, Systems, and Signal Processing*, *42*(9), 5559–5578. https://doi.org/10.1007/s00034-023-02374-7

Liang, H., Chen, L., Pan, Y., & Lam, H.-K. (2022). Fuzzy-based robust precision consensus tracking for uncertain networked systems with cooperative–antagonistic interactions. *IEEE Transactions on Fuzzy Systems*, *31*(4), 1362–1376. https://doi.org/10.1109/TFUZZ.2022.3200730

Luo, Z., Ji, J., Wang, B., Shi, D., Ma, H., & Gan, W.-S. (2025). Deep learning-based generative fixed-filter active noise control: Transferability and implementation. *Mechanical Systems and Signal Processing*, *238*, 113207.

Luo, Z., Shi, D., & Gan, W.-S. (2022). A hybrid sfanc-fxnlms algorithm for active noise control based on deep learning. *IEEE Signal Processing Letters*, *29*, 1102–1106. https://doi.org/10.1109/LSP.2022.3169428

Luo, Z., Shi, D., Shen, X., Ji, J., & Gan, W.-S. (2023). Deep generative fixed-filter active noise control. In *Icassp 2023-2023 ieee international conference on acoustics, speech and signal processing (icassp)* (pp. 1–5). IEEE. https://doi.org/10.1109/ICASSP49357.2023.10095205

Maamoun, K. S. A., Wrona, S., Pawelczyk, M., & Karimi, H. R. (2024). Shaping of the frequency response of vibrating plates with openings for vibro-acoustic systems. *Mechanical Systems and Signal Processing*, *218*, 111539. https://doi.org/10.1016/j.ymssp.2024.111539

Maamoun, K. S. A., Wrona, S., Pawelczyk, M., & Karimi, H. R. (2025). Optimizing design of openings in vibrating plates for enhanced vibro-acoustic performance using a genetic algorithm approach. *Journal of Sound and Vibration*, *598*, 118847. https://doi.org/10.1016/j.jsv.2024.118847

Manolakis, D. G., Ingle, V. K., & Kogon, S. M. (2005). Spectral estimation, signal modeling, adaptive filtering, and array processing. *ARTECH HOUSE: Boston, MA, USA*, .

Meng, H., & Chen, S. (2020). A modified adaptive weight-constrained fxLMS algorithm for feedforward active noise control systems. *Applied Acoustics*, *164*, 107227. https://doi.org/10.1016/j.apacoust.2020.107227

Misol, M., Algermissen, S., & Monner, H. P. (2012). Experimental investigation of different active noise control concepts applied to a passenger car equipped with an active windshield. *Journal of Sound and Vibration*, *331*(10), 2209–2219. https://doi.org/10.1016/j.jsv.2012.01.001

Murao, T., Nishimura, M., Sakurama, K., & Nishida, S.-i. (2016). Basic study on active acoustic shielding (improving the method to enlarge the AAS window). *Mechanical Engineering Journal*, *3*(1), 15–00322. https://doi.org/10.1299/mej.15-00322

Napoli, R., & Piroddi, L. (2009). Nonlinear active noise control with NARX models. *IEEE Transactions on Audio, Speech, and Language Processing*, *18*(2), 286–295. https://doi.org/10.1109/TASL.2009.2025798

Shi, D., Gan, W.-S., Lam, B., & Wen, S. (2020). Feedforward selective fixed-filter active noise control: Algorithm and implementation. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *28*, 1479–1492. https://doi.org/10.1109/TASLP.2020.2989582

Shi, D., Lam, B., Ooi, K., Shen, X., & Gan, W.-S. (2022). Selective fixed-filter active noise control based on convolutional neural network. *Signal Processing*, *190*, 108317.

Sivasakthi, D. A., Sathiyaraj, A., & Devendiran, R. (2024). Hybridrobustnet: Enhancing detection of hybrid attacks in iot networks through advanced learning approach. *Cluster Computing*, *27*(4), 5005–5019.

Snyder, S. D., & Tanaka, N. (1995). Active control of vibration using a neural network. *IEEE Transactions on Neural Networks*, *6*(4), 819–828. https://doi.org/10.1109/72.392246

Sun, J., Zhang, H., Wang, Y., & Sun, S. (2020). Fault-tolerant control for stochastic switched IT2 fuzzy uncertain time-delayed nonlinear systems. *IEEE Transactions on Cybernetics*, *52*(2), 1335–1346. https://doi.org/10.1109/TCYB.2020.2997348

Tan, K., & Wang, D. (2019). Learning complex spectral mapping with gated convolutional recurrent networks for monaural speech enhancement. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *28*, 380–390. https://doi.org/10.1109/TASLP.2019.2955276

Tao, J., Wang, S., Qiu, X., & Pan, J. (2016). Performance of an independent planar virtual sound barrier at the opening of a rectangular enclosure. *Applied Acoustics*, *105*, 215–223. https://doi.org/10.1016/j.apacoust.2015.12.019

Tobias, O. J., & Seara, R. (2006). On the LMS algorithm with constant and variable leakage factor in a nonlinear environment. *IEEE Transactions on Signal Processing*, *54*(9), 3448–3458. https://doi.org/10.1109/TSP.2006.879274

Tokhi, M. O., & Wood, R. (1997). Active noise control using radial basis function networks. *Control Engineering Practice*, *5*(9), 1311–1322. https://doi.org/10.1016/S0967-0661(97)84370-5

Vetterli, M., & Le Gall, D. (1989). Perfect reconstruction FIR filter banks: Some properties and factorizations. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *37*(7), 1057–1071. https://doi.org/10.1109/29.32283

Wang, S., Yu, J., Qiu, X., Pawelczyk, M., Shaid, A., & Wang, L. (2017). Active sound radiation control with secondary sources at the edge of the opening. *Applied Acoustics*, *117*, 173–179. https://doi.org/10.1038/s41598-017-13546-2

World Health Organization (2018). Environmental Noise Guidelines for the European Region: Executive Summary. World Health Organization, Regional Office for Europe.

Wrona, S., Pawelczyk, M., & Cheng, L. (2022). Sound transmission through a panel with a hybrid active and semi-active control system. *Journal of Sound and Vibration*, *536*, 117172. https://doi.org/10.1016/j.jsv.2022.117172

Wu, L., He, H., & Qiu, X. (2010). An active impulsive noise control algorithm with logarithmic transformation. *IEEE Transactions on Audio, Speech, and Language Processing*, *19*(4), 1041–1044. https://doi.org/10.1109/TASL.2010.2061227

Zhang, H., & Wang, D. (2021). Deep ANC: A deep learning approach to active noise control. *Neural Networks*, *141*, 1–10. https://doi.org/10.1016/j.neunet.2021.03.037

Zhang, Q.-Z., Gan, W.-S., & Zhou, Y.-l. (2006). Adaptive recurrent fuzzy neural networks for active noise control. *Journal of Sound and Vibration*, *296*(4–5), 935–948. https://doi.org/10.1016/j.jsv.2006.03.020

Zhou, Y., Zhang, Q., & Yin, Y. (2015). Active control of impulsive noise with symmetric α-stable distribution based on an improved step-size normalized adaptive algorithm. *Mechanical Systems and Signal Processing*, *56*, 320–339. https://doi.org/10.1016/j.ymssp.2014.10.002

Transfer Learning-Enhanced DDPG for Adaptive Active Noise Control in Vehicle Interiors

Anonymous Author(s)
Affiliation
Address
email

Abstract

This paper presents an adaptive approach to active noise control (ANC) inside a vehicle cabin by combining deep learning and reinforcement learning techniques. Specifically, a deep neural network-based algorithm, FxMDNoiseNet, is first employed to model and control acoustic noise. To enhance adaptability in dynamic environments, a reinforcement learning algorithm based on Deep Deterministic Policy Gradient (DDPG) is integrated into the control framework. The proposed method leverages transfer learning by initializing the reinforcement learning model with the pre-trained weights of the FxMDNoiseNet. This strategy accelerates convergence and improves overall performance by exploiting prior knowledge acquired during supervised training. The hybrid framework enables the system to adapt to varying noise conditions within the vehicle interior. Experimental results demonstrate improved noise attenuation and robustness, highlighting the effectiveness of combining deep learning, reinforcement learning, and transfer learning for intelligent acoustic control systems.

1 Introduction

Interior noise reduction remains a key challenge in the automotive industry, as it directly impacts passenger comfort and perceived vehicle quality Nelson and Elliott [1991], Aboutiman et al. [2025a]. The acoustic environment inside a vehicle cabin is particularly complex, combining narrowband tonal components generated by the engine with broadband noise originating from road–tire interaction and aerodynamic sources Samarasinghe et al. [2016], Aboutiman et al. [2025b]. Engine-induced noise is especially critical, as periodic excitations related to combustion and rotating components produce strong low-frequency tonal harmonics. These components, typically below 500 Hz, are further amplified by cabin resonances, making them difficult to attenuate using conventional passive noise control techniques.

Active Noise Control (ANC) has therefore emerged as an effective solution for low-frequency noise mitigation, relying on destructive interference to cancel unwanted sound Fahy and Walker [2018]. However, its application in vehicle cabins remains challenging due to the nonlinear, time-varying, and strongly coupled nature of the underlying acoustic system. Engine operating conditions continuously modify the spectral characteristics of the noise, while nonlinearities in actuators and acoustic paths further degrade the performance of traditional adaptive filtering methods such as FxLMS.

To address these limitations, learning-based approaches have been increasingly investigated. Deep learning methods have demonstrated strong capabilities in modeling complex and nonlinear acoustic dynamics, enabling improved noise reduction performance in challenging environments Zhang and Wang [2021], Cha et al. [2023], Aboutiman et al. [2026]. More recently, reinforcement learning (RL) has been introduced to enhance adaptability by allowing control policies to be learned directly through interaction with the environment Luo et al. [2024], Cao et al. [2023], Feng et al. [2023]. In particular,

Submitted to 40th Conference on Neural Information Processing Systems (NeurIPS 2026). Do not distribute.

actor–critic methods such as Deep Deterministic Policy Gradient (DDPG) have shown promising results for continuous control tasks, including active noise control Ryu et al. [2024]. However, most existing studies remain limited to simplified setups or simulation environments, and their applicability to realistic in-vehicle conditions is still largely unexplored.

From a machine learning perspective, in-vehicle ANC presents several key challenges. First, the control system must operate under strongly time-varying and partially unknown environments, mainly due to variations in the secondary acoustic path, making robustness and adaptability more critical than optimal performance under fixed conditions. Second, the disturbance signals are highly non-stationary, combining engine, road, and aerodynamic noise, which requires consistent performance across diverse acoustic scenarios. Finally, real-time deployment imposes strict constraints on computational complexity and memory, limiting the applicability of large-scale models and emphasizing the need for efficient learning-based control strategies.

In this work, we propose a transfer learning-enhanced DDPG for adaptive active noise control framework tailored for vehicle interiors. The approach combines a modified deep learning architecture, referred to as MDNoiseNet, with a DDPG-based reinforcement learning strategy. A filtered-x structure is incorporated to account for the secondary path, resulting in the FxMDNoiseNet framework. To improve training efficiency and convergence, transfer learning is employed by initializing the RL policy with pre-trained FxMDNoiseNet weights. The proposed method is evaluated in offline simulations using real-world in-vehicle noise data, demonstrating its ability to adapt to complex, time-varying acoustic conditions.

2 Problem Formulation and Methodology

The proposed hybrid deep learning and reinforcement learning framework for feedback active noise control (ANC), combining the FxMDNoiseNet algorithm with a Deep Deterministic Policy Gradient (DDPG) strategy to enable adaptive control.

FxMDNoiseNet-based Control. The FxMDNoiseNet algorithm is a deep learning-based feedback control approach designed to generate an anti-noise signal directly from the estimated disturbance. The block diagram is shown in Fig. 1a. In this framework, $d(n)$ denotes the primary disturbance, $a(n)$ the anti-noise generated through the secondary path, $y(n)$ the control signal, $\hat{d}(n)$ the estimated disturbance, and $e(n)$ the residual error. $S(n)$ and $\hat{S}(n)$ represent the true and estimated secondary paths, respectively. At each discrete-time instant n, the input to the network is a sequence of L past samples of the estimated disturbance:

$$\hat{\mathbf{d}}_L(n) = \big(\hat{d}(n-L+1), \ldots, \hat{d}(n)\big)^T, \tag{1}$$

and the control signal is generated as

$$y(n) = \mathrm{FxMDNoiseNet}\big(\hat{\mathbf{d}}_L(n)\big). \tag{2}$$

The anti-noise is obtained through the secondary path:

$$a(n) = S(n) * y(n), \tag{3}$$

and the residual error signal is computed as

$$e(n) = d(n) + a(n). \tag{4}$$

The estimated disturbance is recursively updated as

$$\hat{d}(n) = e(n) + \hat{S}(n) * y(n), \tag{5}$$

which allows the network to account for the secondary path effect and better adapt to the system dynamics.

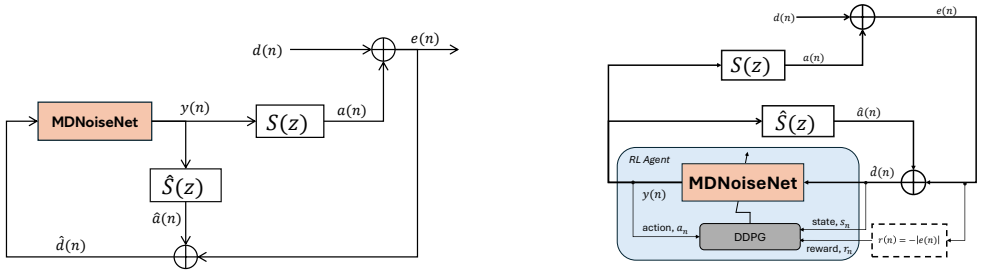
DDPG-based Adaptive Extension. To enhance adaptability in dynamic acoustic environments, the FxMDNoiseNet framework is extended using a DDPG-based reinforcement learning approach. In this formulation, the ANC system is modeled as the environment, while the controller acts as an agent. At each time step n, the agent observes the state $\mathbf{s}_n$, corresponding to the estimated disturbance signal, and outputs a control action:

$$y(n) = \pi_\phi(\mathbf{s}_n), \tag{6}$$

where π_ϕ denotes the actor network parameterized by ϕ. The control signal is propagated through the secondary path to generate the anti-noise, and the residual error $e(n)$ is used to define the reward:

$$r(n) = -|e(n)|. \tag{7}$$

The agent is trained to maximize the cumulative reward, thereby minimizing the residual noise. By leveraging transfer learning, the actor network is initialized using the pre-trained weights of FxMDNoiseNet, which accelerates convergence and improves control performance. Figure 1b illustrates the overall FxMDNoiseNet-DDPG control framework.



(a) (b)

Figure 1: (a) In the feedback ANC configuration, FxMDNoiseNet takes the estimated disturbance $\hat{d}(n)$ as input and outputs the control signal $y(n)$, which is filtered through the secondary path $S(n)$ to yield the anti-noise signal $a(n)$; the residual error $e(n)$, obtained by subtracting $a(n)$ from the actual disturbance $d(n)$, reflects the degree of noise attenuation achieved by the controller. (b) Within the FxMDNoiseNet-DDPG framework, at each discrete step n, the agent receives the state s_n, defined as the estimated disturbance signal, and the actor network maps this state to a control action a_n; the resulting residual error is then leveraged to derive the reward signal r_n, which serves as a performance metric to drive and refine the reinforcement learning optimization.

The secondary path models and primary in-vehicle noise were obtained experimentally in a sedan equipped with a four-cylinder gasoline engine. All signals were recorded at a sampling rate of 3 kHz. Measurements were conducted under three cabin configurations: an empty vehicle, a vehicle with one occupant, and a vehicle with two occupants. These scenarios were selected to capture the effect of varying interior acoustic loading conditions on the secondary path dynamics. Primary noise data were collected under real-world driving conditions on public roads in Germany, including both highway (Autobahn) and urban traffic environments.

3 MDNoiseNet model

3.1 MDNoiseNet architecture

Motivated by Cha et al. [2023], the proposed Modified DNoiseNet (MDNoiseNet) architecture is designed to effectively capture multi-scale temporal features and long-term dependencies from raw time-domain acoustic signals for active noise control. As summarized in Table 1, the model combines causal and dilated convolutional residual blocks with skip connections and a recurrent module, enabling efficient and structured temporal feature extraction. The input to the network consists of a sequence of L past samples of the estimated disturbance signal, denoted by $\hat{\mathbf{d}}_L(n)$. In this case $L = 100$. An initial causal one-dimensional convolution (kernel size 2) followed by a Tanh activation produces a base representation that serves both as the main feature stream and as a reference

for subsequent residual learning. The core of the architecture is composed of a cascade of Atrous Scaled Convolution (ASC) modules with progressively increasing dilation factors. Each module applies a one-dimensional dilated convolution followed by a SeLU activation and a pointwise (1×1) convolution. This design enables an efficient expansion of the receptive field without increasing the model complexity. Residual connections are introduced between successive modules to facilitate gradient propagation and stabilize training, while skip connections are extracted at each stage and progressively aggregated to preserve multi-scale temporal information. Specifically, the ASC modules employ increasing intermediate channel sizes and dilation factors to capture hierarchical temporal patterns at multiple scales. This progressive design allows the network to jointly model short-term fluctuations and long-range dependencies inherent in acoustic signals. Following the convolutional feature extraction stage, the aggregated skip features are combined with the final convolutional output and regularized using a dropout layer. The resulting representation is then reshaped and fed into a two-layer recurrent neural network (RNN), which explicitly models temporal dependencies that may not be fully captured by convolutional operations alone. Finally, the last hidden state of the RNN is passed through two fully connected layers to produce the output control signal. The first layer projects the representation into a latent space with a ReLU activation, while the second layer maps it to a scalar prediction.

Table 1: Detailed architecture of MDNoiseNet.

Block	**Operation**	**Output**	**Connections**
Input	Estimated disturbance $\hat{\mathbf{d}}_L(n)$	x_0	–
Initial Conv	Conv1D (1→1, k=2) + Tanh	x_1	Base signal
	Atrous Scaled Convolutional Blocks		
ASC-1	Dilated Conv (d_1) + SeLU + PW Conv	x_2, s_1	Residual: $x_2 + x_1$; Skip$_1$: s_1
ASC-2	Dilated Conv (d_2) + SeLU + PW Conv	x_3, s_2	Residual: $x_3 + x_2$; Skip$_2$: $s_2 + s_1$
ASC-3	Dilated Conv (d_3) + SeLU + PW Conv	x_4, s_3	Residual: $x_4 + x_3$; Skip$_3$: $s_3 + s_2$
ASC-4	Dilated Conv (d_4) + SeLU + PW Conv	x_5	Residual: $x_5 + x_4$
	Temporal Modeling and Prediction		
Aggregation	Dropout (0.1) + skip aggregation	z	$z = x_5 + s_3$
RNN	2-layer RNN (hidden H)	h	Temporal modeling
FC	Linear ($H \rightarrow 18$) + ReLU	u	–
Output	Linear ($18 \rightarrow 1$)	$y(n)$	–

3.2 Training strategy

The dataset used for the training consists of roughly two hours of in-car noise captured under real driving conditions, from which 10,000 one-second segments were extracted for training. To improve model robustness, synthetic noise was injected into a portion of these samples. A separate five-minute recording, kept entirely out of the training pipeline, was reserved for testing to guarantee an unbiased assessment on unseen data. MDNoiseNet was trained on the 10,000 acoustic segments, each normalized to $[-1, 1]$, with an 80/20 train-validation split and mini-batches of 500 samples. The Adam optimizer was employed with an initial learning rate of 1×10^{-5}, decayed by a factor of 0.5 every 10 epochs to stabilize convergence. Weights were drawn from a zero-mean Gaussian distribution with unit standard deviation, biases were set to zero, and training ran for 100 epochs. At every time step, the model processes a sliding window of the last 100 samples of the estimated disturbance signal. The training objective combines three terms into a single loss function:

$$\mathcal{L} = \mathrm{MSE}(y_{\text{anti}}, y_{\text{target}}) + \lambda_{\text{out}}\, y_{\text{anti}}^2 + \lambda_{\text{L2}} \sum_{i \in \mathcal{B}} \|b_i\|_2^2, \tag{8}$$

where the first term penalizes the reconstruction error between the predicted anti-noise y_{anti} and the target cancellation signal y_{target}, the second limits excessive control energy, and the third applies L2 regularization over the bias parameters $\mathcal{B}$ to curb overfitting. The regularization coefficients are fixed at $\lambda_{\text{out}} = 1 \times 10^{-4}$ and $\lambda_{\text{L2}} = 1 \times 10^{-5}$. Throughout training, an ideal estimation of the secondary path was assumed, using the impulse response measured in the unoccupied vehicle configuration. The training was performed on a machine equipped with an NVIDIA GeForce RTX 4070 GPU (CUDA available), with 67 GB of system RAM.

4 Reinforcement Learning Framework and Training Strategy

4.1 Deep Deterministic Policy Gradient method

Deep Deterministic Policy Gradient (DDPG) Lillicrap et al. [2015] is an off-policy actor–critic reinforcement learning algorithm designed for continuous control problems. It combines the advantages of deterministic policy gradients with deep function approximation and experience replay. DDPG consists of a deterministic actor $\pi_\phi : \mathcal{S} \to \mathcal{A}$, which maps the observed state s_n to a continuous action $a_n = \pi_\phi(s_n)$, and a single critic network $Q_\theta(s_n, a_n)$, which estimates the corresponding action–value function.

To encourage exploration during training, Gaussian noise is commonly added to the actor output. In contrast, this work employs a structured exploration noise tailored to the characteristics of the problem, as detailed in Section 4.2.

$$a_n = \pi_\phi(s_n) + N_n, \tag{9}$$

where $N_n \sim \mathcal{N}(0, \sigma)$ is optionally clipped to satisfy action bounds.

The critic is trained by minimizing the Bellman error using target networks for stability. The target Q-value is defined as:

$$y_n = r_n + \gamma(1 - d_n) Q_{\theta'}(s_{n+1}, \pi_{\phi'}(s_{n+1})), \tag{10}$$

where γ is the discount factor, d_t is the episode termination flag, and θ' and ϕ' denote the parameters of the target critic and actor networks, respectively.

The critic loss is given by:

$$\mathcal{L}_Q = \mathbb{E}\big[(Q_\theta(s_n, a_n) - y_n)^2\big] + \lambda_Q \|\theta\|_2^2, \tag{11}$$

where λ_Q is a regularization coefficient enforcing weight decay on the critic parameters.

The actor is updated by maximizing the expected return estimated by the critic. The objective function is formulated as:

$$\mathcal{L}_\pi = -\mathbb{E}_{s_n \sim \mathcal{D}} \left[Q_\theta(s_n, \pi_\phi(s_n)) \right] + \lambda_\pi \|\phi\|_2^2, \tag{12}$$

where $\mathcal{D}$ denotes the replay buffer and λ_π controls the strength of L2 regularization on the actor parameters.

To stabilize learning, both actor and critic target networks are updated using Polyak averaging:

$$\theta' \leftarrow \tau\theta + (1 - \tau)\theta', \quad \phi' \leftarrow \tau\phi + (1 - \tau)\phi', \tag{13}$$

where $\tau \ll 1$ controls the update rate.

The actor and critic networks are implemented as deep neural architectures tailored to the active noise control task. The actor produces continuous control actions from the observed state, while the critic evaluates state–action pairs. The overall training process is illustrated in Figure 2.

4.2 Training strategy

The training strategy is designed to address the specific constraints of active noise control with real-world engine noise. Beyond the standard Deep Deterministic Policy Gradient (DDPG) framework,

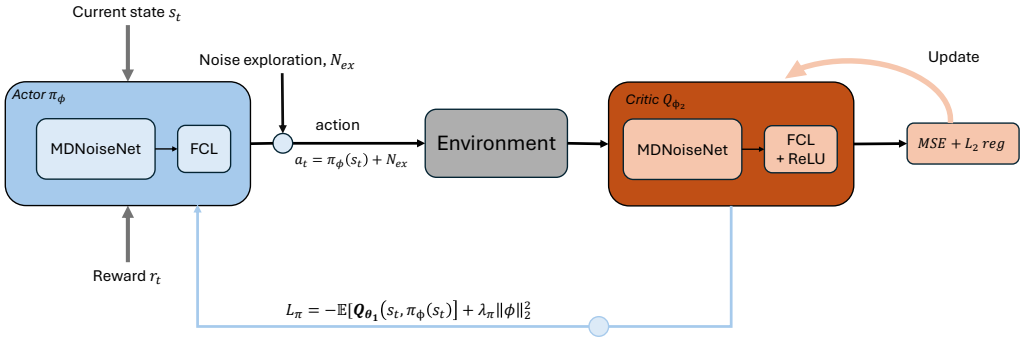


Figure 2: Actor and critic parameter updates.

domain-informed mechanisms are introduced to improve stability, sample efficiency, and physical consistency during learning.

A key challenge in continuous ANC control is exploration. Standard Gaussian noise may excite irrelevant frequency components and degrade stability. To address this, a frequency-adaptive exploration strategy is used. At each time step, the dominant frequency of the residual error signal is estimated via FFT, and the exploration frequency is sampled as

$$f_{\text{ex}} = f_{\text{dom}} + \mathcal{U}(-\Delta f, \Delta f), \tag{14}$$

where Δf defines the exploration bandwidth.

The exploration noise added to the deterministic policy output $a_n = \pi_\phi(s_n)$ is a sinusoidal perturbation with random phase:

$$\boldsymbol{N}_{\text{ex},n} = \rho_t \begin{bmatrix} \cos(2\pi f_{\text{ex}} n + \varphi_n) \\ \sin(2\pi f_{\text{ex}} n + \varphi_n) \end{bmatrix}, \tag{15}$$

with $\varphi_n \sim \mathcal{U}(-\pi, \pi)$ and $\rho_n = \rho_{\text{ex}} \hat{A}_e$, where $\hat{A}_e$ is an estimate of the residual error amplitude.

A warm-up phase is used at the beginning of training, during which actions are sampled uniformly to populate the replay buffer with diverse transitions before policy-driven exploration begins.

The reward is defined as the negative absolute residual error: $r_n = -|e_n|$, which directly encourages noise attenuation. The reward evolution during training is shown in Figure 3.



Figure 3: Evolution of the reward during DDPG training.

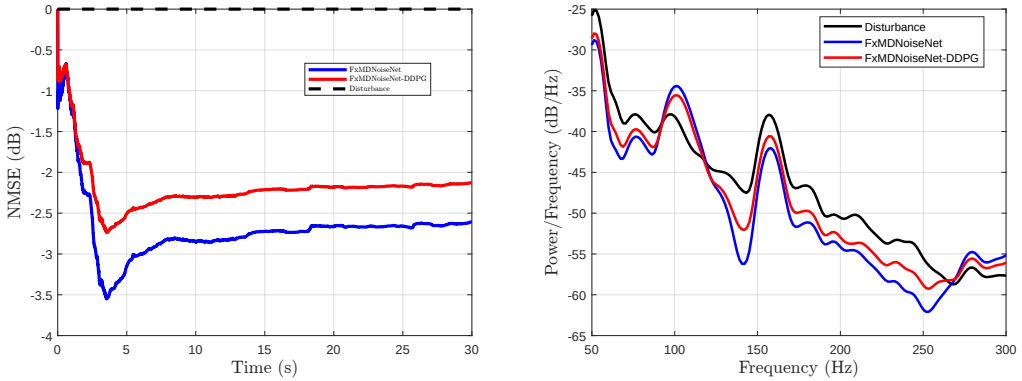
The DDPG training process relies on a set of carefully selected hyperparameters. The actor and critic networks are optimized using learning rates of 2×10^{-5} and 1×10^{-4}, respectively. A discount

factor of $\gamma = 0.2$ is used, together with a soft target update coefficient $\tau = 5 \times 10^{-4}$. The replay buffer size is set to 10^6 transitions, and mini-batches of size 512 are sampled during training. An ℓ_2 regularization term with a weight of 10^{-4} is applied to improve generalization and stability. Training begins with a warm-up phase of 60,000 steps, during which the replay buffer is populated before policy updates are applied. The total number of training steps is set to 500,000. To improve robustness, two different secondary paths (empty vehicle and vehicle with one occupant) are used during training, with the actual secondary path randomly selected at each episode.

5 Results

5.1 Performance with a perfect estimation of the secondary path

In this section, the results obtained with FxMDNoiseNet and FxMDNoiseNet-DDPG are presented. Figure 4a shows the evolution of the Normalized Mean Squared Error (NMSE) over time, while Figure 4b presents the Power Spectral Density (PSD) of both error signals. These results are obtained under the assumption of a perfect estimation of the secondary path.



(a) (b)

Figure 4: Comparison of (a) NMSE over time and (b) the PSD obtained with FxMDNoiseNet (blue) and FxMDNoiseNet-DDPG (red) for a combined engine and road noise.

Figure 4a shows that, when the secondary path is perfectly estimated, FxMDNoiseNet achieves better noise reduction performance than FxMDNoiseNet-DDPG. It is important to note that these results are obtained in the same environment used during training, namely with the secondary path corresponding to the empty vehicle configuration. This suggests that exposing the DDPG-based approach to multiple and varying environments during training may slightly degrade its performance when evaluated in the original training conditions. Figure 4b further confirms this observation in the frequency domain, where FxMDNoiseNet consistently provides better noise attenuation across most frequency components.

5.2 Performance with the secondary path in linear time-variant conditions

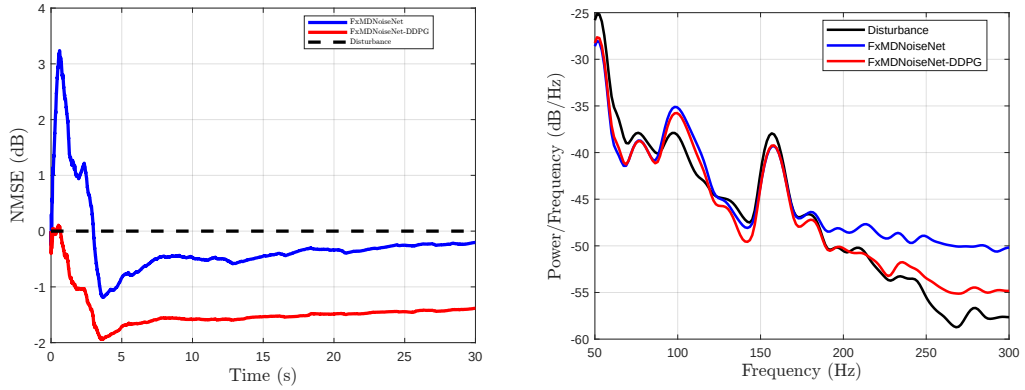
To evaluate robustness under non-stationary conditions, a linear time-varying (LTV) behavior is introduced into the secondary path. At each time step, the true secondary path impulse response is randomly modified using a circular shift operation:

$$\mathbf{s}_n = \mathcal{T}_{\Delta_n}(\mathbf{s}), \quad \Delta_n \in \{-5, 0, 5\}, \tag{16}$$

which corresponds to a 5-sample delay, a 5-sample advance, or no shift, respectively.

In contrast, the estimated secondary path used within the filtered-x structure remains fixed throughout the experiment, resulting in a mismatch between the true and modeled dynamics. This configuration forces the controller to operate under realistic modeling uncertainties and time-varying conditions, providing a meaningful assessment of robustness.

Figure 5a shows that under LTV conditions, FxMDNoiseNet initially exhibits a significant increase in error at the beginning of the signal, during which FxMDNoiseNet-DDPG progressively adapts to



(a) (b)

Figure 5: Comparison of (a) NMSE over time and (b) PSD obtained FxMDNoiseNet (blue), and FxMDNoiseNet-DDPG (red) for a combined engine and road noise scenario.

the changing environment. This highlights that when the control environment is time-varying, the DDPG-based approach enables online adaptation to the new conditions, whereas FxMDNoiseNet, with fixed weights, achieves noise reduction only after a certain adaptation period (approximately 4 seconds in our case). Figure 5b further shows that in the frequency range between 50 Hz and 300 Hz, FxMDNoiseNet-DDPG provides better noise attenuation across most frequency components.

6 Conclusion

This paper proposes a transfer learning-enhanced DDPG framework to improve the performance of deep learning-based active noise control systems, with a particular focus on FxMDNoiseNet. The proposed FxMDNoiseNet-DDPG approach leverages the strong representation learning capabilities of deep neural networks together with the adaptability of reinforcement learning, enabling more effective control in complex and time-varying acoustic environments.

The experimental results demonstrate that the proposed method enhances FxMDNoiseNet, particularly in terms of adaptability. Under perfect secondary path estimation, the original FxMDNoiseNet achieves the best overall noise attenuation while effectively limiting high-frequency amplification. However, under linear time-varying conditions, the proposed approach exhibits superior robustness, especially in the low-frequency range, which is critical for active noise control applications.

Several directions can be explored in future work to further improve the proposed framework. First, the current FxMDNoiseNet operates in a sample-by-sample fashion, producing one control action at each time step. Extending this formulation to a block-based scheme, where multiple samples are processed jointly, could improve temporal consistency and enhance stability. Second, although DDPG is adopted in this study, other deterministic reinforcement learning algorithms, such as TD3 or D4PG, could be investigated to potentially improve performance and training efficiency. Finally, implementing the proposed approach in a real-time experimental setup would be an important step toward practical deployment, enabling validation under real acoustic conditions and hardware constraints.

References

Alkahf Aboutiman, Zulfi Rachman, Tin Oberman, Francesco Aletta, Jian Kang, Hamid Reza Karimi, and Francesco Ripamonti. Subjective perception analysis of active noise control algorithms in an encapsulated structure: An experimental study. *Applied Acoustics*, 239:110823, 2025a. ISSN 0003-682X. doi: https://doi.org/10.1016/j.apacoust.2025.110823. URL https://www.sciencedirect.com/science/article/pii/S0003682X25002956.

Alkahf Aboutiman, Reza Shams, Hamid Reza Karimi, Francesco Ripamonti, and Marek Pawełczyk. Active noise control in encapsulated structures with non-minimum phase characteristics using a kalman filter approach. *Journal of Sound and Vibration*, 615:119187, 2025b. ISSN 0022-460X. doi: https://doi.org/10.1016/j.jsv.2025.119187. URL https://www.sciencedirect.com/science/article/pii/S0022460X25002615.

Alkahf Aboutiman, Khaled Said Ahmed Maamoun, Hamid Reza Karimi, and Francesco Ripamonti. Hybrid deep learning–based active noise control for encapsulated structures with openings. *Expert Systems with Applications*, page 131247, 2026. doi: https://doi.org/10.1016/j.eswa.2026.131247. URL https://www.sciencedirect.com/science/article/pii/S0003682X25002956.

Zhong Cao, Kun Jiang, Weitao Zhou, Shaobing Xu, Huei Peng, and Diange Yang. Continuous improvement of self-driving cars using dynamic confidence-aware reinforcement learning. *Nature Machine Intelligence*, 5(2):145–158, 2023.

Young-Jin Cha, Alireza Mostafavi, and Sukhpreet S Benipal. Dnoisenet: Deep learning-based feedback active noise control in various noisy environments. *Engineering Applications of Artificial Intelligence*, 121:105971, 2023. doi: https://doi.org/10.1016/j.engappai.2023.105971.

Frank Fahy and John Walker. *Advanced applications in acoustics, noise and vibration*. CRC Press, 2018.

Shuo Feng, Haowei Sun, Xintao Yan, Haojie Zhu, Zhengxia Zou, Shengyin Shen, and Henry X Liu. Dense reinforcement learning for safety validation of autonomous vehicles. *Nature*, 615(7953): 620–627, 2023.

Timothy P Lillicrap, Jonathan J Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*, 2015.

Zhengding Luo, Haozhe Ma, Dongyuan Shi, and Woon-Seng Gan. Gfanc-rl: Reinforcement learning-based generative fixed-filter active noise control. *Neural Networks*, 180:106687, 2024.

Philip Arthur Nelson and Stephen J Elliott. *Active control of sound*. Academic press, 1991.

Seokhoon Ryu, Jihea Lim, and Young-Sup Lee. Narrowband active noise control with ddpg based on reinforcement learning. *International Journal of Automotive Technology*, 25(6):1389–1397, 2024.

Prasanga N Samarasinghe, Wen Zhang, and Thushara D Abhayapala. Recent advances in active noise control inside automobile cabins: Toward quieter cars. *IEEE Signal Processing Magazine*, 33(6): 61–73, 2016.

Hao Zhang and DeLiang Wang. Deep anc: A deep learning approach to active noise control. *Neural Networks*, 141:1–10, 2021. doi: https://doi.org/10.1016/j.neunet.2021.03.037.

Deep transfer reinforcement learning for active noise control in vehicle interiors

Alkahf Aboutiman[1,2*], Jan Kirchhof[2], Hamid Reza Karimi[1], Francesco Ripamonti[1]

[1*]Department of Mechanical Engineering, Politecnico di Milano, Via Privata Giuseppe La Masa, 1, Milano, 20156, Italy.

[2]Müller-BBM Active Sound Technology GmbH, Helmut-A.-Müller-Straße 1 - 5, Munich, 82152, Germany.

*Corresponding author(s). E-mail(s): alkahfahmat.aboutiman@polimi.it;
Contributing authors: jan.kirchhof@mbbm-ast.com; hamidreza.karimi@polimi.it;
francesco.ripamonti@polimi.it;

Abstract

This paper investigates an active noise control (ANC) framework enhanced through deep transfer reinforcement learning. The proposed approach is built upon FxMDNoiseNet, a neural network-based controller operating within a filtered-x architecture. To improve adaptability, a deterministic reinforcement learning strategy called Twin Delayed Deep Deterministic Policy Gradient (TD3) is integrated to fine-tune the control policy and enhance performance under both linear time-invariant (LTI) and linear time-varying (LTV) conditions, including scenarios with secondary acoustic path mismatch. The proposed method is evaluated in a realistic automotive ANC scenario, where experimentally measured impulse responses from a real vehicle are used to model the primary and secondary acoustic paths. This allows the simulation environment to closely reflect real-world in-vehicle acoustic dynamics. Experimental results show that the reinforcement learning-enhanced model achieves improved noise attenuation compared to both the standalone deep learning controller and the classical FxNLMS algorithm.

Keywords: Active noise control, Deep transfer reinforcement learning, Vehicule interior

1 Introduction

Minimization of interior cabin noise has been a major research topic in the automotive industry over the past decades, driven by the increasing demand for ride comfort, which strongly influences passenger perception and purchasing decisions [1, 2]. The acoustic environment inside a vehicle cabin is particularly challenging, as interior noise typically consists of a combination of narrowband tonal components related to engine rotation and broadband noise mainly induced by road–tire interaction and aerodynamic excitations [3, 4]. Among these contributions, engine-related noise plays a dominant role. From a mechanical perspective, engine operation generates periodic excitation forces associated with combustion processes and rotating components, which are transmitted through multiple structural paths to the vehicle body and cabin enclosure. These excitations induce structural vibrations that are subsequently radiated as airborne sound inside the cabin, producing narrowband tonal components directly linked to the engine rotational speed and its harmonics [3, 5]. In addition, structure-borne road noise, characterized by mechanical vibrations transmitted through the tires and vehicle structure, contributes significantly to the overall cabin noise, with most of its acoustic energy concentrated below 1 kHz [6]. A common characteristic of these dominant noise sources is their strong low-frequency content, typically below 500 Hz, where acoustic resonances of the enclosed cabin further amplify sound pressure levels. In this frequency range, conventional passive noise control techniques,

such as sound insulation, damping, and absorption treatments, are generally less effective and often conflict with the automotive industry's requirements for lightweight and cost-efficient vehicle designs. As a consequence, Active Noise Control (ANC) has emerged as a promising alternative for in-vehicle noise mitigation. Based on the principle of destructive interference, ANC systems employ secondary sound sources to generate anti-noise signals with the same amplitude and opposite phase as the disturbing sound field, thereby reducing the residual noise measured by error microphones inside the cabin [5].

Despite the effectiveness of active noise control for attenuating low-frequency disturbances, its practical deployment in vehicle interiors remains particularly challenging due to the complex, time-varying, and nonlinear nature of automotive acoustic systems. In real driving conditions, the presence of an operating engine introduces non-stationary excitation mechanisms whose spectral characteristics evolve continuously with engine speed, load, and driving maneuvers. Unlike simplified laboratory ANC setups relying on stationary acoustic sources, such as noise boxes, an operating engine gives rise to multiple coupled mechanical and acoustic excitations whose characteristics vary over time, thereby substantially increasing the complexity of the control problem. These conditions lead to time-varying harmonic structures and strongly coupled vibro-acoustic propagation paths within the vehicle cabin, which significantly complicate the design of robust and adaptive control strategies. Moreover, several inherent nonlinearities further degrade the performance of conventional ANC algorithms. These include the nonlinear electro-mechanical behavior of loudspeakers operating at high sound pressure levels, nonlinear distortions introduced by power amplifiers, and amplitude-dependent variations of the secondary acoustic paths caused by structural vibrations and changing boundary conditions inside the cabin [7]. The combined effect of non-stationary excitations and system nonlinearities often results in slow convergence, reduced robustness, and potential instability when using classical adaptive filtering techniques, such as the filtered-x least mean squares (FxLMS) algorithm, which relies on linear and time-invariant system assumptions.

Nonlinearities and non-stationary behaviors are intrinsic characteristics of in-vehicle ANC systems, arising from factors such as engine speed variations, actuator dynamics, and time-varying acoustic paths. These properties significantly limit the effectiveness of conventional linear control strategies and have motivated the exploration of learning-based methods capable of adapting to complex system dynamics. Early research efforts relied on shallow neural models, particularly multilayer perceptrons [8], trained online with adaptive filtered-x backpropagation to approximate nonlinear mappings between reference signals and control actions. Subsequent developments improved convergence and computational efficiency [9, 10] and explored alternative neural architectures such as functional link networks [11, 12], radial basis function networks [13], fuzzy neural networks [14], and recurrent neural networks [15]. While these studies demonstrated the potential of neural-based controllers to compensate for nonlinear distortions, most results were obtained under controlled laboratory conditions or simplified experimental setups. More recently, deep learning approaches have gained prominence due to their ability to model highly nonlinear and time-varying dynamics without explicit linear assumptions. [16] proposed a convolutional recurrent neural network (CRN) for ANC, effectively capturing both spectral and temporal dependencies in complex noise fields. [17] introduced DNoisenet, a deep learning-based feedback ANC framework capable of handling nonlinearities across various noisy environments. [18] applied deep learning-based ANC in construction site scenarios, showing strong generalization under unpredictable and highly variable acoustic conditions. Additionally, hybrid learning-based approaches in encapsulated structures with openings have been explored to further enhance adaptive capabilities in dynamic acoustic environments [19]. Despite these advances, practical deployment of deep learning-based ANC in real vehicle cabins remains limited. Key challenges include online adaptation, robustness to rapidly changing engine and road noise, and stability in the presence of nonlinear and time-varying secondary paths. These limitations highlight the need for more autonomous and adaptive learning frameworks capable of reliable operation under real-world automotive conditions, motivating the integration of reinforcement learning and transfer learning strategies for ANC.

Reinforcement learning (RL) provides a promising approach to enhance the adaptability and robustness of deep learning-based active noise control algorithms [20, 21]. By learning directly from interaction with the acoustic environment, RL enables online adaptation and improved noise reduction. However, research in this area remains very limited, especially in realistic automotive settings. For instance, the GFANC-RL framework employs a convolutional neural network trained with the Soft Actor-Critic (SAC) algorithm, removing the need for labeled data and handling non-differentiability issues, while achieving effective noise suppression across different acoustic paths [22]. Similarly, a narrowband ANC study using Deep Deterministic Policy Gradient (DDPG) applies deep RL without a secondary path model, demonstrating robust control and stability under path variations in real-time experiments [23]. Other studies have focused on simulation-based evaluations. In [24], the authors applied a Q-learning method [25] to a narrowband ANC problem,

showing that the approach could maintain robust control stability against sudden changes in the secondary path without relying on a path model. Although these results are promising, they remain preliminary and primarily validated in simplified or simulated environments, highlighting the need for further research in realistic in-vehicle conditions.

In this study, we propose a deep transfer reinforcement learning-based active noise control algorithm, specifically designed to address the challenges of nonlinear and time-varying acoustic environments in vehicle cabins. Our approach builds upon a modified version of the DNoisenet architecture, originally presented in [17]. This modified architecture, referred to as MDNoiseNet (Modified DNoiseNet), is designed to achieve a more effective handling of the secondary path. To this end, a Filtered-x structure is incorporated into the network. The resulting complete active noise control algorithm is denoted FxMDNoiseNet (Filtered-x Modified DNoiseNet). To enhance adaptability and reduce training time, we employ a Deep Transfer Reinforcement Learning (DTRL) strategy. In this framework, the pre-trained weights of FxMDNoisenet are used to initialize the policy network of the reinforcement learning agent. This transfer learning approach provides a baseline for the RL agent, accelerating convergence during the training. Crucially, the proposed FxMDNoisenet-DTRL algorithm is evaluated through offline simulations using real-world data, demonstrating its effectiveness under realistic operating conditions. Experimental results highlight the algorithm's ability to adapt to complex, time-varying engine and road noise, offering a promising and scalable solution for automotive ANC applications. The main contributions of this paper are threefold: (i) the integration of a reinforcement learning strategy into a deep learning-based active noise control framework; (ii) the implementation and validation of this algorithm in a complex acoustic environment, specifically a vehicle interior; and (iii) the development of an ANC algorithm based on a filtered-x structure combined with a deep learning approach.

The remainder of this paper is structured as follows. First, the experimental setup used to compute the impulse responses of the control environment and to construct the dataset for training is presented. Second, the proposed ANC algorithm is introduced, along with its neural network architecture and training procedure. Third, the reinforcement learning strategy and its training process are described. The results obtained with the proposed method are then presented and compared with traditional approaches. Finally, a conclusion summarizes the main contributions of the paper and outlines directions for future work.

2 System Identification and dataset construction

Secondary path models as well as primary noise inside the vehicle cabin were obtained experimentally. The measurements were conducted in a sedan car equipped with a four-cylinder gasoline engine. Figure 1 shows views of the vehicle from both the exterior and the interior.



(a) (b)

Fig. 1: Picture of the experimental environment: (a) interior view and (b) exterior view of the vehicle.

The sampling rate used for all recorded signals was 3 kHz. As secondary path models, impulse responses corresponding to a five-speaker to four-microphone configuration were measured under different passenger occupancy conditions using the Multiple Exponential Sine Sweep Method (MESM) [26]. The system was implemented on a SHARC SC598 DSP, ensuring that the measured impulse responses capture the full closed-loop behavior of a realistic active noise control (ANC) setup. The excitation system consisted of the four door-mounted woofers together with the factory-installed subwoofer located in the trunk of the vehicle. Four

Mikrotek Geffel M380 1/4-inch measurement microphones were used for recording. They were positioned at the four seating locations inside the cabin, corresponding to the driver seat, front passenger seat, and the two rear passenger seats. Measurements were carried out for three different cabin configurations: empty vehicle, vehicle with one occupant, and vehicle with two occupants. These configurations were used to capture the influence of varying interior acoustic loading conditions on the secondary path dynamics. Figure 2 shows the frequency response functions of the measured secondary paths for the three occupancy conditions used in the simulations.

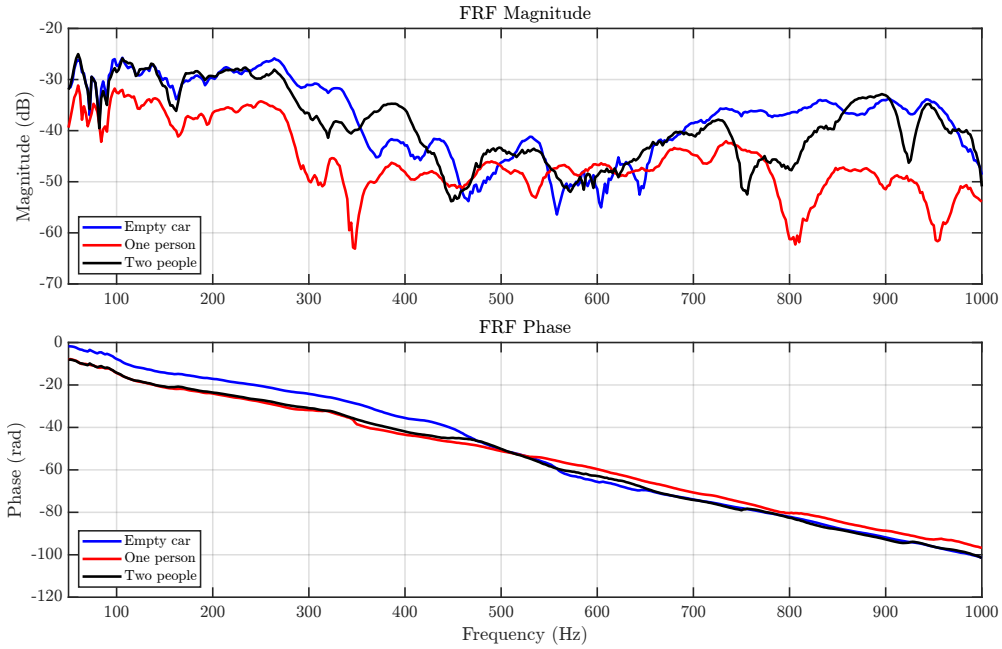


Fig. 2: Used Secondary Transfer Paths: empty car (blue), only driver inside (red), and driver and front right (black) passenger seat occupied

The primary noise measurements were recorded on public German roads, including both Autobahn and inner-city traffic conditions, using the same microphones as those employed for the secondary path measurements. The signals were acquired using a PAK MKII frontend. The resulting dataset consists of approximately two hours of in-vehicle noise recorded during real driving conditions. From this dataset, 10,000 one-second signals were extracted for training purposes. Artificial noise was added to a subset of these signals. In addition, a separate five-minute segment of recorded driving noise was deliberately excluded from the training process and reserved exclusively for evaluation of the proposed algorithm, ensuring an unbiased assessment of its performance on unseen data.

3 Filtered-x Modified DNoiseNet methodology

3.1 Filtered-x Modified DNoiseNet algorithm

The proposed FxMDNoisenet algorithm is a deep learning based feedback control strategy designed to generate an anti-noise signal directly from the estimated disturbance.

The block diagram of the FxMDNoisenet algorithm is shown in Figure 3. In the figure, $d(n)$ denotes the primary disturbance, $a(n)$ the anti-noise produced by the secondary path, $y(n)$ the control signal generated by FxMDNoisenet, $\hat{d}(n)$ the estimated disturbance, and $e(n)$ the residual error. $S(n)$ and $\hat{S}(n)$ represent the impulse response of the secondary path and its estimation, respectively.

At each discrete-time instant n, the input to MDNoisenet consists of a sequence of L past samples of the estimated disturbance:

$$\hat{\mathbf{d}}_L(n) = \big(\hat{d}(n-L+1), \hat{d}(n-L+2), \ldots, \hat{d}(n)\big)^T, \tag{1}$$

and the control signal is generated as

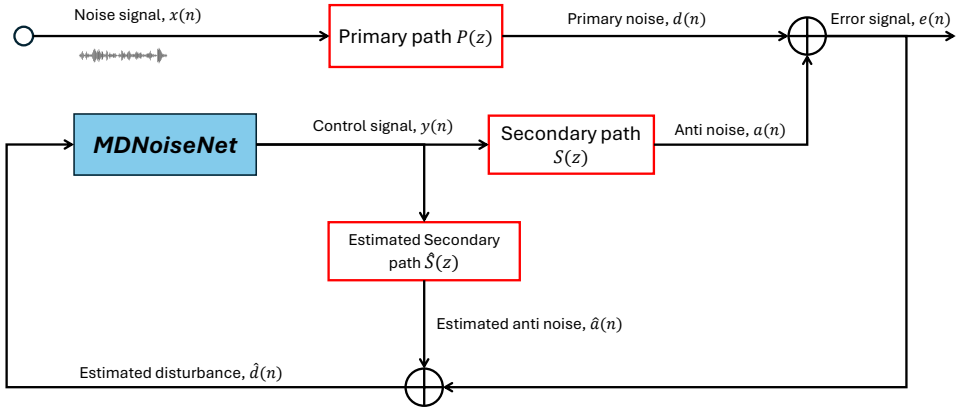


Fig. 3: FxMDNoisenet for feedback ANC. The network takes the estimated disturbance $\hat{d}(n)$ as input and generates the control signal $y(n)$, which is filtered through the secondary path $S(n)$ to produce the anti-noise $a(n)$. The residual $e(n)$ is obtained by subtracting $a(n)$ from the disturbance $d(n)$.

$$y(n) = \text{MDNoisenet}\big(\hat{\mathbf{d}}_L(n)\big). \tag{2}$$

The anti-noise produced by the secondary path is

$$a(n) = S(n) * y(n), \tag{3}$$

and the residual error signal is computed as

$$e(n) = d(n) - a(n), \tag{4}$$

where $*$ denotes linear convolution. The estimated disturbance used as input to the network is recursively updated according to

$$\hat{d}(n) = e(n) + \hat{S}(n) * y(n), \tag{5}$$

where $\hat{S}(n)$ is the estimated secondary path. This recursive strategy ensures that FxMDNoisenet observes a realistic error signal that incorporates both the residual noise and the effect of the estimated secondary path, allowing the network to adapt to the dynamic behavior of the system. Table 1 gives a step-by-step presentation of the algortihm.

Table 1: Step-by-step description of the FxMDNoisenet algorithm with corresponding equations.

Step Description	Corresponding Equation
Construct input sequence of estimated disturbance	$\hat{\mathbf{d}}_L(n) = (\hat{d}(n-L+1), \ldots, \hat{d}(n))^T$
Generate control signal with MDNoiseNet	$y(n) = \text{MDNoisenet}(\hat{\mathbf{d}}_L(n))$
Propagate control signal through the secondary path	$a(n) = S(n) * y(n)$
Compute residual error	$e(n) = d(n) - a(n)$
Update estimated disturbance for next step	$\hat{d}(n+1) = e(n) + \hat{S}(n) * y(n)$

3.2 MDNoiseNet architecture and computational complexity

Motivated by [17], the proposed Modified DNoiseNet (MDNoiseNet) is designed to operate directly on raw time-domain acoustic signals and to accurately capture their temporal patterns under real-time constraints. The architecture integrates causal and atrous scaled convolutions together with residual and skip connections, enabling efficient multi-scale temporal feature extraction while preserving strict causality.

The input to the network consists of a single-channel one-dimensional acoustic waveform. An initial causal one-dimensional convolution with one input channel, one output channel, and a kernel size of 2

is applied to extract low-level temporal features from the input signal. Causal zero-padding is performed exclusively on past samples. This convolutional layer is followed by a Tanh activation function, which introduces nonlinearity and stabilizes the signal amplitude. The resulting feature map is used both as the main feature stream and as an initial skip signal for subsequent processing stages.

The core of the architecture is composed of a stack of Atrous Scaled Convolution (ASC) modules. Each ASC module employs a one-dimensional atrous convolution with increasing dilation factors to progressively enlarge the temporal receptive field without increasing the number of learnable parameters. This design allows the network to capture both short-term and long-term temporal dependencies inherent in acoustic signals. Each atrous convolution is followed by a SeLU activation function to promote self-normalizing behavior. A pointwise (1×1) convolution is then applied to reduce the channel dimensionality and to generate both a residual output and a skip contribution. Residual connections add the module output to its input to facilitate gradient flow, while skip connections accumulate multi-scale temporal information across successive ASC modules.

The first ASC module expands the feature representation using 10 intermediate channels and applies an atrous convolution with a dilation factor of 1, followed by a pointwise convolution that projects the features back to a single channel. The second and third ASC modules employ 20 and 30 intermediate channels, respectively, with atrous convolutions using dilation factors of 2 and 4. This progressive increase in dilation enables hierarchical temporal feature extraction at multiple time scales. The final ASC module uses 40 intermediate channels and applies an atrous convolution with a dilation factor of 8, followed by a SeLU activation and a pointwise convolution. A residual addition is then performed to further refine the extracted temporal features.

Following the convolutional feature extraction stage, the aggregated skip features are regularized using a dropout layer with a dropout rate of 10 percent. The resulting feature sequence is then reshaped and provided as input to a two-layer recurrent neural network (RNN) with a hidden size of 2. The RNN explicitly models temporal dependencies that may not be fully captured by convolutional operations alone, providing an additional level of temporal context integration. Only the output corresponding to the final time step is retained, yielding a fixed-dimensional temporal representation.

Finally, the extracted representation is passed through two fully connected layers. The first fully connected layer maps the RNN output to an 18-dimensional latent space and is followed by a ReLU activation function. The second linear layer projects this representation to a single scalar output. This output corresponds to the control signal.

By combining causal atrous scaled convolutions, hierarchical residual learning, and recurrent temporal modeling, the Modified DNoiseNet effectively captures multi-scale temporal patterns directly in the time domain. This architecture is particularly well suited for active noise control applications, where adaptation to the temporal and spectral characteristics of acoustic signals is essential. The detailed architecture of the proposed network is presented in Figure 4.

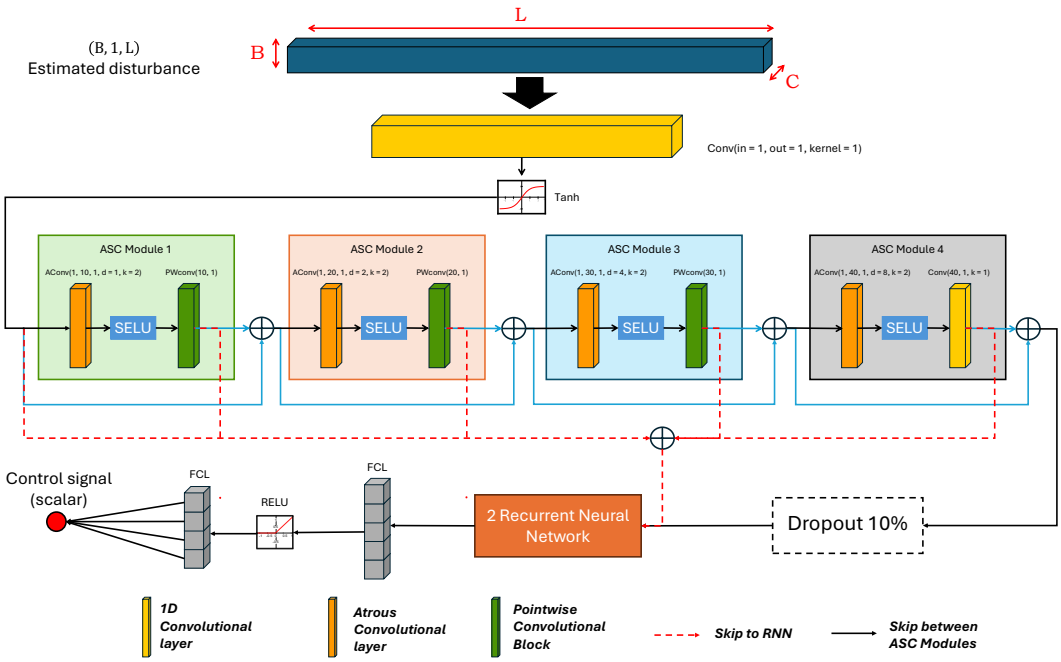


Fig. 4: Architecture of the proposed MDNoiseNet. The network is composed of 1D causal, pointwise, and atrous convolutions to extract effective features from small raw time-domain acoustic signals. The hyperparameters of each layer are specified in the form (input channels, output channels, kernel size, dilation). For layers without dilation, the dilation factor is denoted as d when applicable.

To assess the computational complexity of the proposed framework, we analyze MDNoiseNet in a single-channel configuration, which is used as a controlled benchmark setting for architectural evaluation. The model contains a total of 565 trainable parameters. Despite its compact size, the convolutional layers account for most of the computational load. Using standard analysis, the total number of Multiply–Accumulate operations (MACs) required for a single forward pass is measured as 49.71 KMAC. Table 2 summarizes the parameter count and MAC cost of the network. From a multi-channel perspective, the computational cost is expected to scale approximately linearly with the number of microphone–loudspeaker pairs, as each channel requires an independent control signal estimation. However, the per-channel complexity of MDNoiseNet remains unchanged, and parameter sharing strategies could be considered in future work to mitigate computational growth in large-scale ANC deployments.

Table 2: Computational complexity of the proposed MDNoiseNet model.

Component	Parameters	MACs per forward pass
MDNoiseNet	565	49.71 K

3.3 Training strategy

The training of the proposed MDNoiseNet was performed using 10,000 one-second acoustic signals, originally from the dataset presented in Section 2, normalized to the range $[-1, 1]$. The dataset was split into 80% training and 20% validation sets, and training was performed using mini-batches of 500 signals. The network was optimized using the Adam optimizer with a learning rate of 1×10^{-5}. A learning rate scheduler was applied with a step size of 10 epochs and a decay factor of 0.5. The network weights were initialized using a zero-mean Gaussian distribution with a standard deviation of 1, while all bias parameters were initialized to zero. Training was conducted for 100 epochs. At each time step, the model receives as input a sequence of the last 100 samples of the estimated disturbance. The loss function is defined as the sum of three terms.

The first term is the mean squared error (MSE) between the predicted anti-noise signal and the target noise signal. The second term is an energy regularization applied to the control output to penalize excessive control effort. The third term is an L2 regularization applied only to the bias parameters of the network:

$$\mathcal{L} = \mathrm{MSE}(y_{\mathrm{anti}}, y_{\mathrm{target}}) + \lambda_{\mathrm{out}}\, y_{\mathrm{anti}}^2 + \lambda_{\mathrm{L2}} \sum_{i \in \mathcal{B}} \|b_i\|_2^2, \tag{6}$$

The coefficients λout and λ_{L2} are regularization weights that control the contribution of the output energy and bias regularization terms, respectively, with $\lambda_{\mathrm{out}} = 1 \times 10^{-4}$ and $\lambda_{\mathrm{L2}} = 1 \times 10^{-5}$. y_{anti} denotes the anti-noise signal predicted by the network, and y_{target} is the desired noise signal to be canceled. The set $\mathcal{B}$ denotes the collection of all bias parameters in the network, and b_i represents the i-th bias term. The regularization term $\sum_{i \in \mathcal{B}} |b_i|_2^2$ enforces small bias values to improve generalization and prevent overfitting. During training, the estimation of the secondary path was perfect and the secondary path used was the impulse response of the empty car. Figure 5 shows the resulting training and validation loss curves.

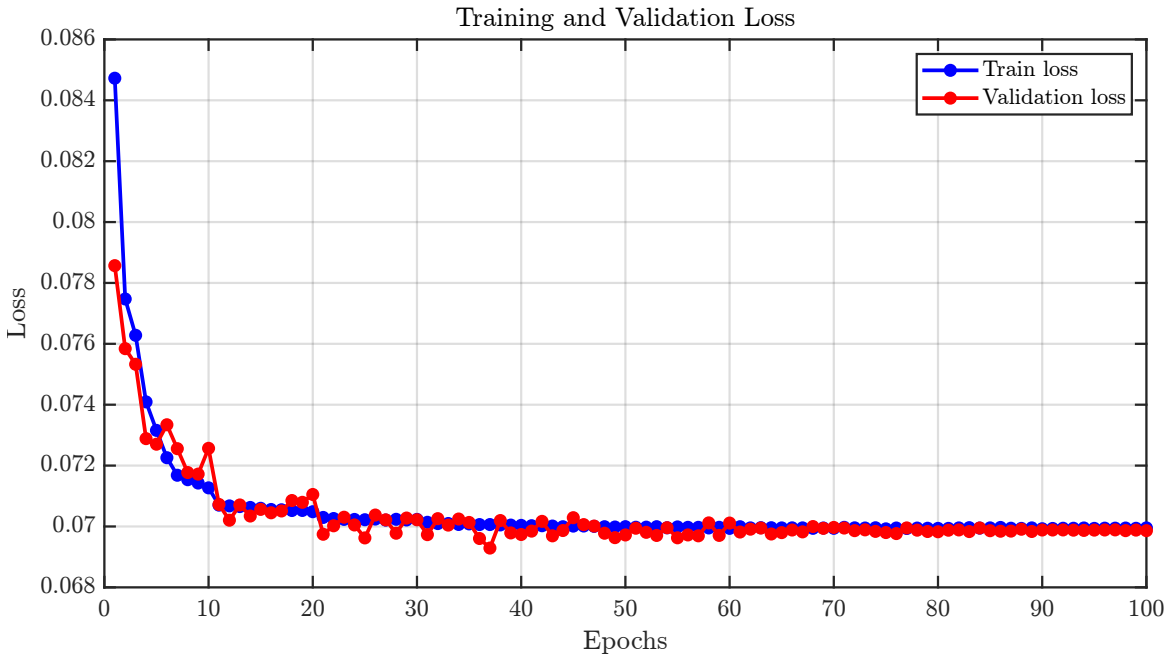


Fig. 5: Training and validation loss of FxMDNoiseNet.

4 Filtered-x Modified DNoiseNet Deep Transfer Reinforcement Learning methodology

4.1 Deep transfer reinforcement learning methodology

Deep Transfer Reinforcement Learning (DTRL) combines the representation learning capability of deep neural networks with the decision-making framework of reinforcement learning. The core idea of DTRL is to leverage knowledge acquired from a pre-trained deep learning model and transfer it to a reinforcement learning (RL) agent, thereby accelerating convergence, improving stability, and reducing the amount of exploration required during training. In this work, the weights of the deep neural network presented in the Section 3 are used as an initialization point for the reinforcement learning model. Instead of training the RL agent from scratch, the pre-trained model provides a structured and informative feature representation of the acoustic environment. This transfer of knowledge allows the RL agent to focus on policy optimization rather than low-level feature extraction, which is beneficial in complex and non-stationary noise control scenarios.

Reinforcement learning is formulated as a sequential decision-making problem in which an agent interacts with an environment. At each discrete time step n, the agent observes a state s_n, selects an action a_n according to a policy $\pi(a_n|s_n)$, and receives a scalar reward r_n that reflects the quality of the chosen action. The environment then transitions to a new state s_{n+1}, and the process is repeated. The objective of the

RL agent is to learn an optimal policy that maximizes the expected cumulative reward over time. Figure 6 illustrates the overall Deep Transfer Reinforcement Learning framework adopted in this study, highlighting the transfer of knowledge from the pre-trained deep learning model to the reinforcement learning agent.

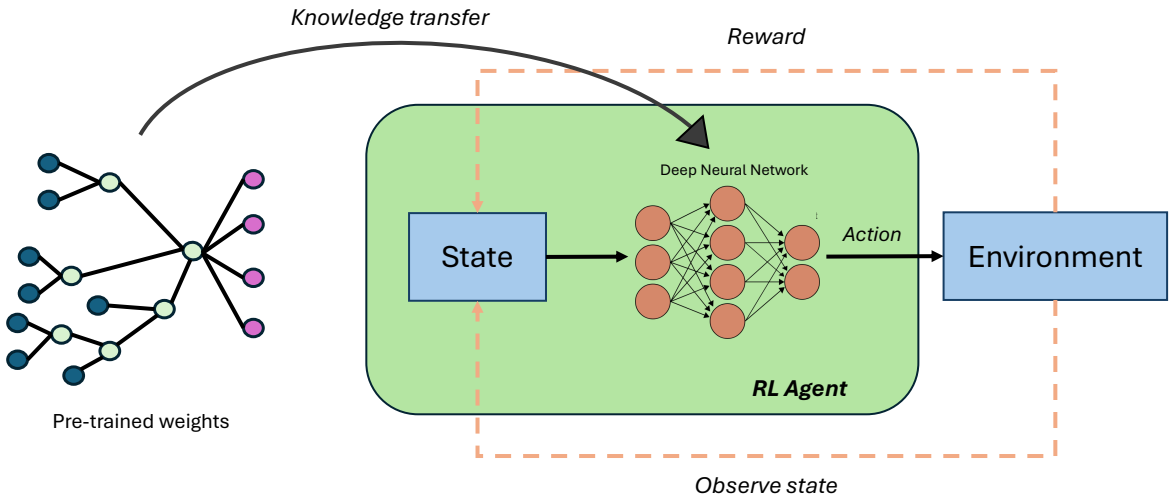


Fig. 6: Scheme of the deep transfer reinforcement learning concept.

4.2 The proposed FxMDNoiseNet-DTRL method

In this approach, the noise cancellation system acts as the environment, while the FxMDNoisenet-DTRL agent and its policy updates are considered the RL agent. The agent outputs a control action $y(n) = \pi_\phi(\mathbf{s}_n)$, which is applied through the secondary path to generate the anti-noise. The residual error is observed and fed back to the agent, and the reward at each step is defined as the negative absolute value of this error, $r(n) = -|e(n)|$. Over training, the RL algorithm optimizes the agent's policy to maximize cumulative reward, effectively reducing the residual noise. Once training converges, the learned policy can be used for noise control. This section provides a detailed overview of the design and application of the FxMDNoisenet-DTRL algorithm. Figure 7 presents the control application using the algorithm.

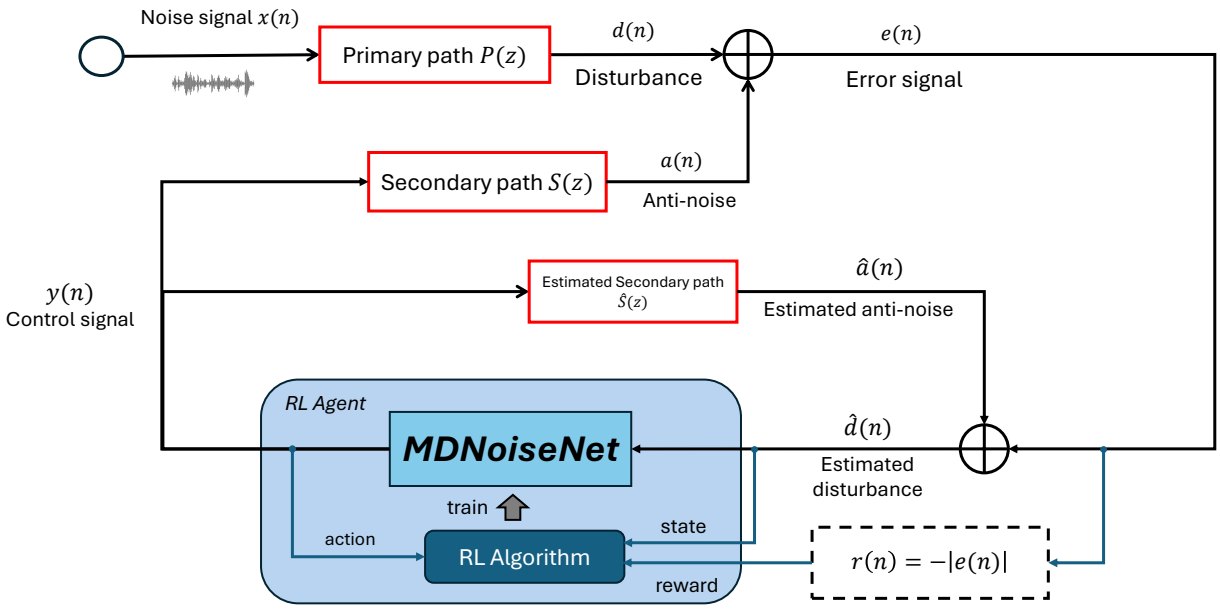


Fig. 7: Block diagram of the FxMDNoiseNet-DTRL framework. At time step n, the agent observes the state s_n, corresponding to the estimated disturbance signal. The actor network generates the control action a_n, which is the control signal. The residual error signal is then used to compute the reward r_n, which evaluates the control performance and guides the learning process.

4.2.1 RL Algorithm: Twin Delayed Deep Deterministic Policy Gradient

Twin Delayed Deep Deterministic Policy Gradient (TD3) [27] is an off-policy actor–critic reinforcement learning algorithm for continuous control. TD3 extends Deep Deterministic Policy Gradient (DDPG) [28] by explicitly addressing two fundamental issues observed in deterministic actor–critic methods: overestimation bias in value functions and high variance in policy updates.

Overestimation bias originates from errors in function approximation when learning action–value functions. In value-based reinforcement learning, the maximization over noisy value estimates leads to systematically overestimated targets [29]. Although less obvious in actor–critic methods, [27] demonstrated that deterministic policy gradient algorithms also suffer from overestimation, since the policy is directly optimized with respect to an approximate critic. Errors in the critic can therefore propagate through the policy update, creating a feedback loop in which suboptimal actions are consistently reinforced.

To mitigate this effect, TD3 introduces Clipped Double Q-learning, which maintains two independent critic networks Q_{θ_1} and Q_{θ_2}. Instead of relying on a single value estimate, the target Q-value is computed using the minimum of the two target critics:

$$y_n = r_n + \gamma(1 - d_n) \min_{i=1,2} Q_{\theta_i'}(s_{n+1}, \tilde{a}_{n+1}), \tag{7}$$

where γ is the discount factor and d_n indicates episode termination. By taking the minimum, this target update prevents the introduction of additional overestimation. Although it may induce mild underestimation, this behavior is preferable, as underestimated actions are not amplified through the policy update [27, 30].

Each critic network is trained by minimizing the Bellman error, augmented with an ℓ_2 regularization term on the network parameters:

$$\mathcal{L}_Q = \mathbb{E}\Big[(Q_{\theta_1}(s_n, a_n) - y_n)^2 + (Q_{\theta_2}(s_n, a_n) - y_n)^2\Big] + \lambda_Q \left(\|\theta_1\|_2^2 + \|\theta_2\|_2^2\right), \tag{8}$$

where λ_Q controls the strength of the regularization.

In addition to overestimation, deterministic policies are sensitive to narrow peaks in the learned value function, which often arise from approximation errors. To reduce the variance of the critic target and improve robustness, TD3 employs target policy smoothing. The target action is perturbed with clipped Gaussian

noise:

$$\tilde{a}_{n+1} = \pi_{\phi'}(s_{n+1}) + \epsilon, \quad \epsilon \sim \text{clip}(\mathcal{N}(0, \sigma), -c, c). \tag{9}$$

This regularization encourages smoother value estimates and is conceptually related to Expected SARSA [25], assigning higher value to actions that are robust to small perturbations.

In addition to target policy smoothing, TD3 relies on exploration noise to encourage sufficient exploration of the action space during training. Specifically, the action executed in the environment is perturbed with an exploration noise term $N_{\text{ex},n}$:

$$a_n = \pi_\phi(s_n) + N_{\text{ex},n}, \tag{10}$$

where $N_{\text{ex},t}$ is commonly sampled from a Gaussian distribution and optionally clipped to remain within the action bounds of the environment. In this work, the exploration noise is specifically designed for the ANC task and is described in Section 4.2.2. This exploration noise allows the agent to discover better policies by visiting a wider range of state–action pairs, which is particularly important in continuous control tasks [27, 28].

TD3 further improves training stability by decoupling value learning from policy updates. While the critic networks are updated at every learning step, the actor network is updated only once every d iterations. This delayed policy update allows the critics to reduce their estimation error before being used to update the policy, leading to more reliable policy gradients. Such two-timescale updates are known to be beneficial for convergence in actor–critic methods [31].

To stabilize temporal-difference learning, TD3 relies on target networks for both the actor and critics, updated using Polyak averaging:

$$\theta_i' \leftarrow \tau\theta_i + (1-\tau)\theta_i', \quad \phi' \leftarrow \tau\phi + (1-\tau)\phi', \tag{11}$$

where $\tau \ll 1$ controls the update rate. Slowly evolving target networks reduce the accumulation of approximation errors and have become a standard stabilization technique in deep reinforcement learning [32].

The actor network $\pi_\phi : \mathcal{S} \to \mathcal{A}$ is a deterministic policy parametrized by ϕ, which maps the observed state s_n to a continuous control action $a_n = \pi_\phi(s_n)$. The actor is optimized to maximize the expected action-value estimated by the first critic Q_{θ_1}, while a ℓ_2 regularization term is introduced to improve training stability. The resulting objective function is given by

$$\mathcal{L}_\pi = -\mathbb{E}_{s_n \sim \mathcal{D}}\left[Q_{\theta_1}(s_n, \pi_\phi(s_n))\right] + \lambda_\pi \|\phi\|_2^2, \tag{12}$$

where $\mathcal{D}$ denotes the replay buffer used to sample mini-batches of past transitions, θ_1 are the parameters of the first critic network, and λ_π controls the strength of the ℓ_2 regularization applied to the actor parameters.

Using a single critic for the policy gradient reduces computational complexity while preserving the benefits of clipped double Q-learning.

The actor and critic networks are implemented as deep neural architectures adapted to the considered control task. The actor maps the state representation to a continuous control action, while each critic jointly processes the state and action to estimate the corresponding Q-value. The detailed architectures of the actor and critic networks are illustrated in Figure 8.

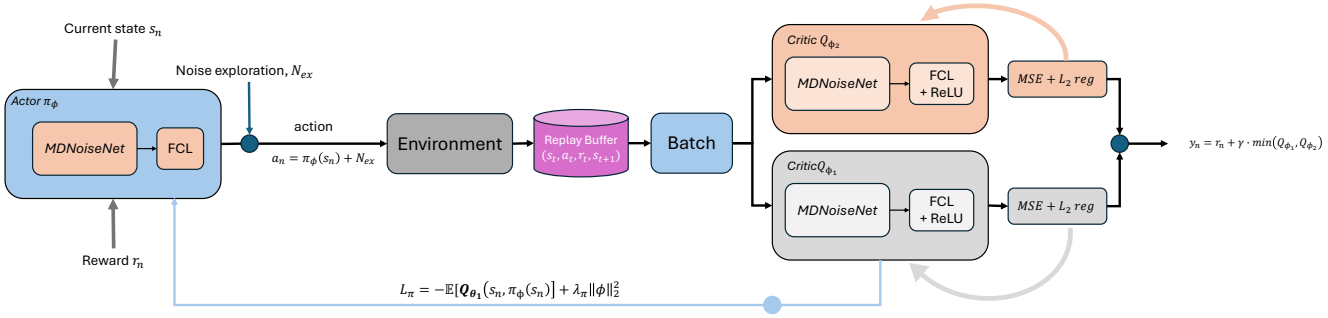


Fig. 8: Actor and critic parameters updates.

4.2.2 Training Strategy

The training strategy is specifically designed to address the characteristics and constraints of active noise control with real-world engine noise. Beyond the standard TD3 framework, several domain-informed mechanisms are introduced to improve stability, sample efficiency, and physical relevance during learning.

Exploration is a critical challenge in continuous control for active noise control (ANC), as standard Gaussian perturbations may excite acoustically irrelevant frequencies and potentially destabilize the control process. To overcome this limitation, a frequency-adaptive exploration strategy is adopted, specifically tailored to the spectral characteristics of engine noise. At each interaction step, the dominant frequency of the current residual error signal, represented in the agent state s_n, is estimated using a fast Fourier transform (FFT). Let f_{dom} denote the frequency corresponding to the maximum spectral amplitude. The exploration frequency f_{ex} is then sampled from a narrow band centered around this dominant component,

$$f_{\mathrm{ex}} = f_{\mathrm{dom}} + \mathcal{U}(-\Delta f, \Delta f), \tag{13}$$

where Δf controls the exploration bandwidth.

The exploratory action perturbation added to the deterministic policy output $a_n = \pi_\phi(s_n)$ is generated as a two-dimensional sinusoidal signal with random phase,

$$\boldsymbol{N}_{\mathrm{ex},n} = \rho_n \begin{bmatrix} \cos(2\pi f_{\mathrm{ex}} n + \varphi_n) \\ \sin(2\pi f_{\mathrm{ex}} n + \varphi_n) \end{bmatrix}, \tag{14}$$

where $\varphi_n \sim \mathcal{U}(-\pi, \pi)$ is resampled at each time step. The exploration gain ρ_n is defined as

$$\rho_n = \rho_{\mathrm{ex}}\, \hat{A}_e, \tag{15}$$

with ρ_{ex} a fixed scaling factor and $\hat{A}_e$ an estimate of the residual error amplitude. This design ensures structured and physically meaningful exploration, consistent with the sinusoidal nature of engine disturbances, while adaptively adjusting both frequency content and magnitude throughout training.

During an initial warm-up period, actions are sampled uniformly from the admissible action space. This phase allows the replay buffer to be populated with sufficiently diverse transitions before policy-driven actions are introduced, reducing early overfitting and improving critic initialization.

The ASC modules are initially frozen by setting their learning rate to zero and disabling gradient updates in the critic networks. The unfreezing process starts at training step $t_{\mathrm{ASC}} = 90{,}000$. From this point, the ASC parameters are progressively activated by enabling gradient computation and linearly increasing their learning rate over a duration of $\Delta t_{\mathrm{ASC}} = 200{,}000$ steps. With a maximum learning rate of $\alpha_{\mathrm{ASC}}^{\max} = 1 \times 10^{-6}$. This progressive unfreezing strategy allows the ASC modules to adapt smoothly to the reinforcement learning objective while preserving the pretrained representations, thereby reducing the risk of catastrophic forgetting.

To further enhance stability, ℓ_2 regularization is applied to both actor and critic parameters during optimization. In addition, gradient norms are clipped to prevent exploding gradients, which can occur due to the strong coupling between the policy and value estimates in deterministic actor–critic methods.

In addition to exploration during environment interaction, structured noise is also applied when computing target actions for critic updates. This target action perturbation follows the same frequency-adaptive design as the exploration noise, encouraging the critics to learn value estimates that are robust to small, physically meaningful deviations in the control signal.

The reward is directly derived from the residual acoustic error measured at the error microphone. At each time step n, the reward is defined as

$$r_n = -|e_n|, \tag{16}$$

where e_n denotes the instantaneous residual error signal. This formulation encourages the agent to minimize the absolute error amplitude, which directly corresponds to maximizing noise attenuation. The variation of this reward throughout training is analyzed in Figure 9.

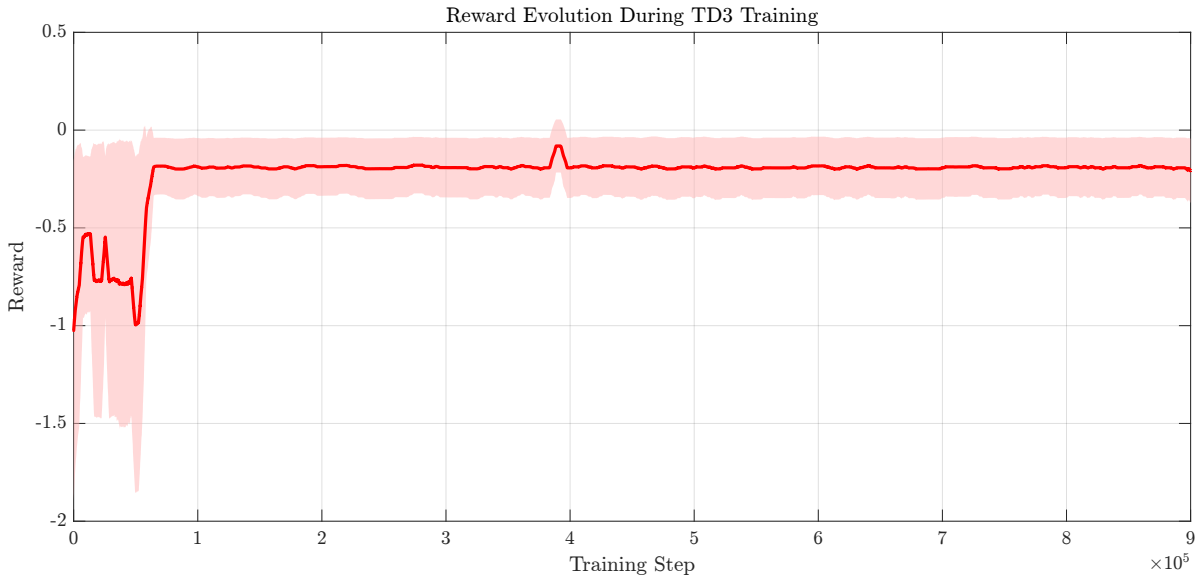


Fig. 9: Reward evolution during the training.

The hyperparameters used for training are summarized in Table 3. Note that during training, two different secondary paths were used: the impulse response of the empty car and that of the car with one occupant. The estimated secondary path corresponded to the impulse response of the empty car. Moreover, during training, the actual secondary path was randomly selected for each episode in order to improve the robustness of the algorithm to variations in the secondary path.

Table 3: TD3 training hyperparameters used for the ANC task.

Hyperparameter	**Value**
Actor learning rate (α_π)	2×10^{-5}
Critic learning rate (α_Q)	1×10^{-4}
Discount factor (γ)	0.2
Soft update coefficient (τ)	0.005
Policy delay	3
Replay buffer size	1×10^6
Batch size	512
ℓ_2 regularization (λ_π, λ_Q)	1×10^{-4}
Warm-up steps	60,000
Total training steps	900,000
Unfreeze start for ASC modules (Critic)	90,000 steps
ASC modules max learning rate ($\alpha_{\text{ASC}}^{\max}$)	1×10^{-6}

5 Results

This section presents a comprehensive comparison between FxNLMS, FxMDNoiseNet and FxMDNoiseNet-DTRL under several challenging acoustic conditions. The objective is to evaluate and contrast their noise reduction performance, adaptability, and robustness in realistic ANC scenarios. To ensure a fair comparison between all algorithms, the initial weights of the FxNLMS were pre-trained. During the application of the control, the estimation of the disturbance was filtered with a low-pass filter at 800 Hz. This technique provides improved stability.

In the first set of experiments, these algorithms are exposed to a combined engine and road noise environment. These tests aim to assess their effectiveness in handling signals with increasing spectral and temporal complexity.

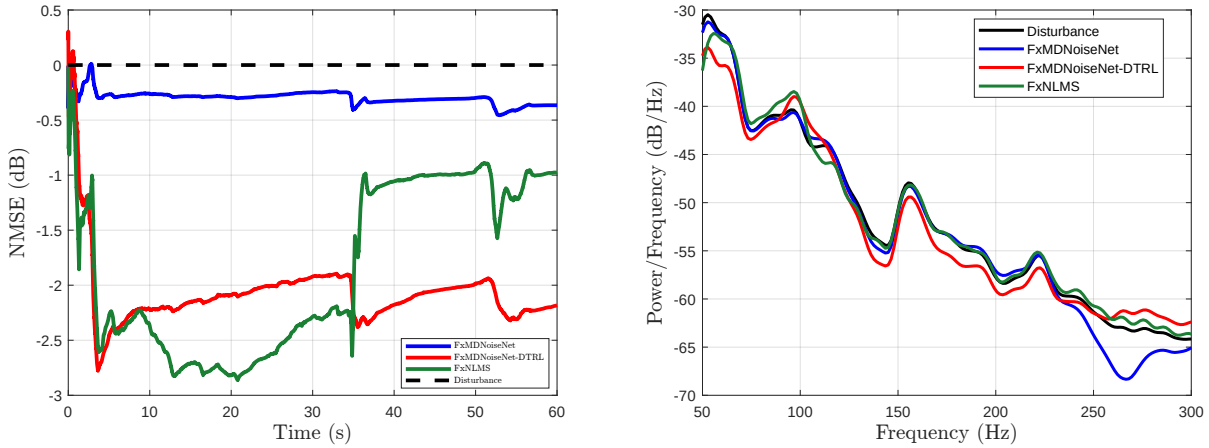
In the second set of experiments, the adaptability of the algorithms is evaluated by introducing a secondary path that was not encountered during training. This scenario reflects practical ANC conditions

where system characteristics may differ from those seen during the learning phase, thereby testing the generalization capability of each approach.

Finally, the robustness of FxMDNoiseNet and FxMDNoiseNet-DTRL is examined by exposing them to a linear time-varying (LTV) secondary path. This experiment evaluates their ability to maintain stable and effective noise attenuation in the presence of dynamic system variations, which commonly occur in real-world applications.

5.1 ANC Performances

In this section, results are obtained under a perfect estimation of the secondary path, the impulse response of the empty car was used. Figure 10a presents the evolution of the Normalized Mean Square Error (NMSE) over time. It can be observed that the best noise reduction over the 60-second signal is achieved by FxMDNoiseNet-DTRL in comptetion with FxNLMS. FxMDNoisenet achieves the highest NMSE over time. During the first 35 seconds, FxNLMS shows a lower NMSE; however, it starts to increase after this period. In contrast, FxMDNoiseNet remains stable throughout the entire signal duration.



(a) (b)

Fig. 10: Comparison of (a) NMSE over time and (b) the PSD obtained with FxNLMS (green), FxMDNoiseNet (blue) and FxMDNoiseNet-DTRL (red) for a combined engine and road noise.

Figures 10b and 11 illustrate the error signals in the frequency domain. In this context, FxMDNoiseNet-DTRL achieves superior attenuation at low frequencies, and thanks to the filtering applied to the estimated disturbance, no amplification is observed at high frequencies. Moreover, compared to the high-frequency amplification reported in [17], the proposed method exhibits reduced amplification. This improvement can be attributed to the penalty imposed on the controller output in equation 6, which regularizes the control effort and limits excessive responses at higher frequencies. The overall noise reduction results further support these observations: FxMDNoiseNet-DTRL achieves 2.2 dB, FxMDNoiseNet 0.4 dB, and FxNLMS 1 dB, respectively.

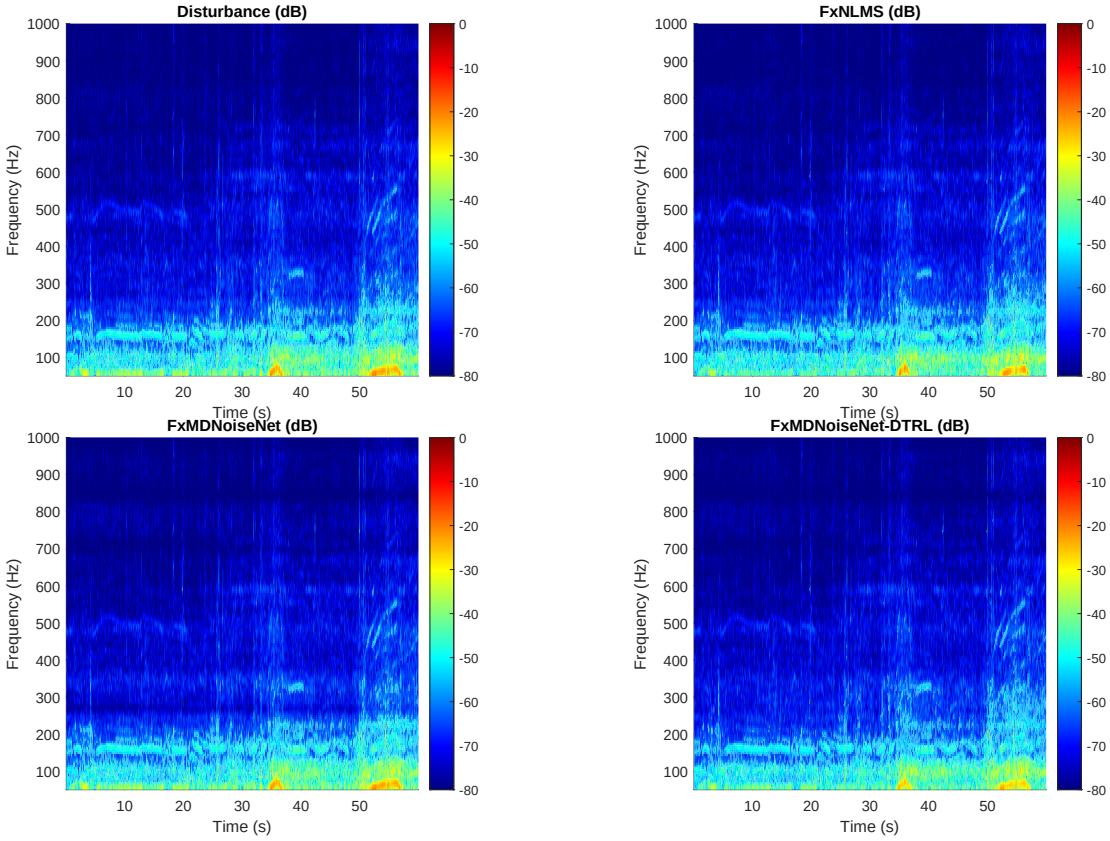
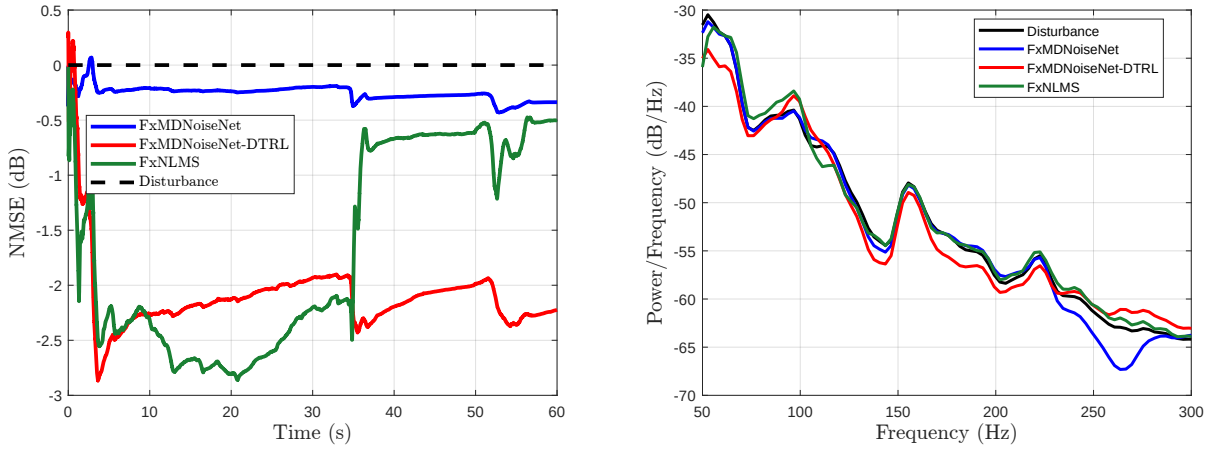


Fig. 11: Comparison of spectrogram of the primary noise, the error signal obtained with FxNLMS, FxMDNoiseNet and FxMDNoiseNet-DTRL for a combined engine and road noise.

5.2 ANC Adaptability

In order to highlight the adaptability of the proposed algorithms, their performance is evaluated under imperfect secondary path estimation. In this study, the impulse response of the empty vehicle is used as the estimated secondary path, while the impulse response of the vehicle with two occupants is used as the true secondary path, which is not seen during training. Figure 12a shows that FxMDNoiseNet continues to provide consistent noise reduction over time, but the level remains low. Thanks to the filtering of the estimated disturbance the control obtained with this algorithm doesn't present unstability. FxNLMS and FxMDNoiseNet-DTRL both achieve meaningful noise reduction; however, FxMDNoiseNet-DTRL provides the best performance.



(a) (b)

Fig. 12: Comparison of (a) NMSE over time and (b) PSD obtained using FxNLMS (green), FxMDNoiseNet (blue), and FxMDNoiseNet-DTRL (red) for a combined engine and road noise scenario. In this case, the impulse response of the empty vehicle is used as an estimation of the secondary path, while the impulse response of the vehicle with two occupants is used as the true secondary path.

Figures 12b and 13 depict the signals in the frequency domain under imperfect secondary path estimation. In this scenario, FxMDNoiseNet-DTRL provides more effective attenuation in the low-frequency range, particularly between 50 and approximately 220 Hz. However, around 260 Hz, FxMDNoiseNet shows slightly better noise reduction performance. Overall, the global noise reduction results emphasize the limited adaptability of FxMDNoiseNet when the secondary path is not perfectly estimated, as evidenced by a degradation of 0.3 dB. In contrast, FxMDNoiseNet-DTRL demonstrates the best performance, achieving 2.2 dB, followed by FxNLMS with 0.5 dB.

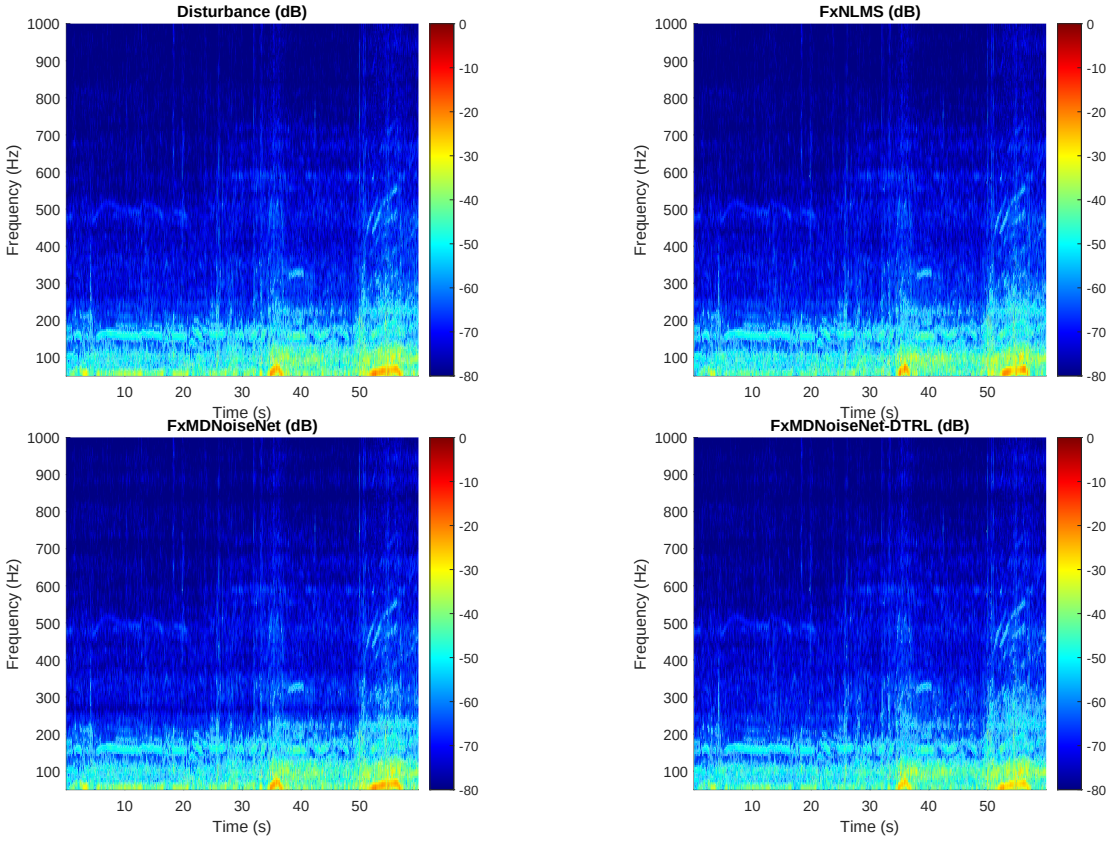


Fig. 13: Comparison of spectrogram of the primary noise, the error signal obtained with FxNLMS, FxMDNoiseNet and FxMDNoiseNet-DTRL for a combined engine and road noise. In this case, the impulse response of the empty vehicle is used as an estimate of the secondary path, while the impulse response of the vehicle with two occupants is used as the true secondary path.

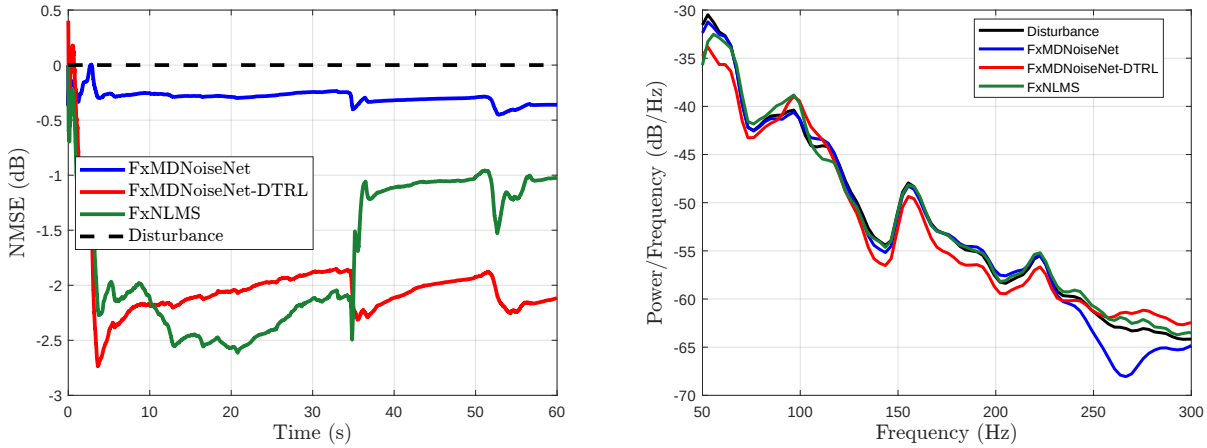
5.3 ANC Robustness

To assess robustness under non-stationary conditions, a linear time-varying (LTV) behavior is introduced in the secondary path. At each time step, the true secondary path impulse response is randomly shifted using a circular shift,

$$\mathbf{s}_n = \mathcal{T}_{\Delta_n}(\mathbf{s}), \quad \Delta_n \in \{-1, 0, 1\}, \tag{17}$$

which corresponds to adding a delay of one sample, advancing it by one sample, or leaving it unchanged at each iteration. The estimated secondary path used in the filtered-x structure remains fixed, creating a mismatch between the true and modeled paths. This setup forces the controller to operate under realistic modeling errors and time variations, providing a practical evaluation of robustness.

Figure 14a shows that, in the time domain, FxMDNoiseNet-DTRL achieves the best noise reduction performance. Similar to the previous section, FxMDNoiseNet provides a fairly consistent noise reduction over time, although it remains low. The most notable noise reductions were achieved by FxMDNoiseNet and FxNLMS.



(a) (b)

Fig. 14: Comparison of (a) NMSE over time and (b) PSD obtained using FxNLMS (green), FxMDNoiseNet (blue), and FxMDNoiseNet-DTRL (red) for a combined engine and road noise scenario. In this case, the impulse response of the empty vehicle is used as an estimate of the secondary path, while the impulse response of the vehicle with two occupants is used as the true secondary path.

Figures 14b and 15 show that, in the frequency domain, FxMDNoiseNet-DTRL achieves the best noise reduction performance in the low-frequency range, particularly between 50 and 250 Hz. This confirms its effectiveness in attenuating the dominant engine noise components, which are mainly concentrated in this band. This trade-off highlights the inherent challenge of achieving robust broadband noise reduction under LTV conditions. In terms of overall performance, FxMDNoiseNet-DTRL achieves the highest global noise reduction with 2.2 dB, outperforming both FxMDNoiseNet with 0.3 dB and FxNLMS with 0.5 dB.

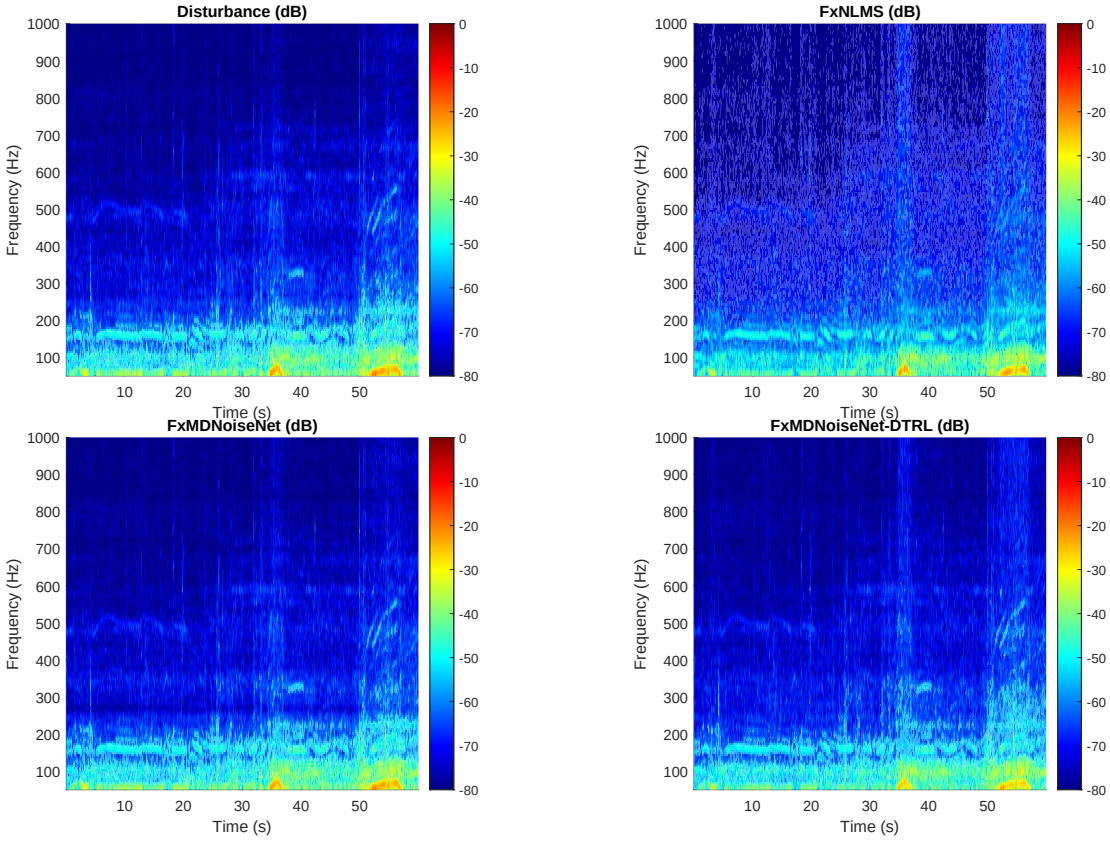


Fig. 15: Comparison of spectrogram of the primary noise, the error signal obtained with FxNLMS, FxMDNoiseNet and FxMDNoiseNet-DTRL for a combined engine and road noise. In this case, the impulse response of the empty vehicle is used as an estimate of the secondary path, while the impulse response of the vehicle with two occupants is used as the true secondary path.

6 Conclusion

This paper introduced a deep transfer reinforcement learning framework to improve the performance of deep learning-based active noise control, specifically FxMDNoiseNet. The proposed FxMDNoiseNet-DTRL approach combines the representation capabilities of neural networks with the adaptability of reinforcement learning, enabling enhanced control in complex acoustic environments.

The experimental results demonstrate that the proposed method consistently outperforms both FxNLMS and the baseline FxMDNoiseNet across different scenarios. Under perfect secondary path estimation, FxMDNoiseNet-DTRL achieves the highest global noise reduction while effectively limiting high-frequency amplification. In the presence of model mismatch, it shows superior adaptability, maintaining stable performance where the baseline model degrades. Furthermore, under linear time-varying conditions, the proposed approach exhibits strong robustness, particularly in the low-frequency range where active noise control is most critical.

Future work could explore several directions to further improve the proposed approach. First, the current FxMDNoiseNet operates in a sample-by-sample manner, generating one control output per time step. Extending the framework to a block-based implementation, where multiple samples are processed simultaneously, could improve stability and better capture temporal dependencies. Second, while the TD3 algorithm was used in this work, alternative deterministic reinforcement learning methods, such as Deep Deterministic Policy Gradient variants like D4PG, could be investigated to potentially achieve improved performance and training efficiency. Finally, a real-time implementation of the proposed method would be an important step toward practical deployment, allowing validation under real acoustic conditions and hardware constraints.

Contributions

A.A. conceived the project, contributed to the theoretical understanding, and developed numerical simulations. J.K. contributed to data acquisition and simulation computation. H.R.K. and F.R. contributed to the supervision of the work. All authors contributed to the experimental design, scientific presentation, discussion, and writing of the manuscript.

Acknowledgments

The work was supported by the project IN-NOVA: Active reduction of noise transmitted into and from enclosures through encapsulated structures, which has received funding from the European Union's Horizon Europe programme under the Marie Sklodowska-Curie grant agreement no. 101073037.

Competing interests statement

The authors declare no competing interest.

References

[1] Nelson, P. A. & Elliott, S. J. *Active control of sound* (Academic press, 1991).

[2] Aboutiman, A. *et al.* Subjective perception analysis of active noise control algorithms in an encapsulated structure: An experimental study. *Applied Acoustics* **239**, 110823 (2025). URL https://www.sciencedirect.com/science/article/pii/S0003682X25002956.

[3] Samarasinghe, P. N., Zhang, W. & Abhayapala, T. D. Recent advances in active noise control inside automobile cabins: Toward quieter cars. *IEEE Signal Processing Magazine* **33**, 61–73 (2016).

[4] Aboutiman, A., Shams, R., Karimi, H. R., Ripamonti, F. & Pawełczyk, M. Active noise control in encapsulated structures with non-minimum phase characteristics using a kalman filter approach. *Journal of Sound and Vibration* **615**, 119187 (2025). URL https://www.sciencedirect.com/science/article/pii/S0022460X25002615.

[5] Fahy, F. & Walker, J. *Advanced applications in acoustics, noise and vibration* (CRC Press, 2018).

[6] Belgacem, W., Berry, A. & Masson, P. Active vibration control on a quarter-car for cancellation of road noise disturbance. *Journal of Sound and Vibration* **331**, 3240–3254 (2012).

[7] Elliot, S. J. Active control of nonlinear systems. *Noise Control Engineering Journal* **49**, 30–53 (2001).

[8] Snyder, S. D. & Tanaka, N. Active control of vibration using a neural network. *IEEE Transactions on Neural Networks* **6**, 819–828 (1995).

[9] Bouchard, M., Paillard, B. & Le Dinh, C. T. Improved training of neural networks for the nonlinear active control of sound and vibration. *IEEE transactions on neural networks* **10**, 391–401 (1999).

[10] Chang, C.-Y. & Luoh, F.-B. Enhancement of active noise control using neural-based filtered-x algorithm. *Journal of sound and vibration* **305**, 348–356 (2007).

[11] Das, D. P. & Panda, G. Active mitigation of nonlinear noise processes using a novel filtered-s lms algorithm. *IEEE Transactions on Speech and Audio Processing* **12**, 313–322 (2004).

[12] Krukowicz, T. Active noise control algorithm based on a neural network and nonlinear input-output system identification model. *Archives of Acoustics* **35**, 191–202 (2010).

[13] Tokhi, M. & Wood, R. Active noise control using radial basis function networks. *Control Engineering Practice* **5**, 1311–1322 (1997).

[14] Zhang, Q.-Z., Gan, W.-S. & Zhou, Y.-l. Adaptive recurrent fuzzy neural networks for active noise control. *Journal of Sound and Vibration* **296**, 935–948 (2006).

[15] Bambang, R. T. Adjoint ekf learning in recurrent neural networks for nonlinear active noise control. *Applied Soft Computing* **8**, 1498–1504 (2008).

[16] Zhang, H. & Wang, D. Deep anc: A deep learning approach to active noise control. *Neural Networks* **141**, 1–10 (2021).

[17] Cha, Y.-J., Mostafavi, A. & Benipal, S. S. Dnoisenet: Deep learning-based feedback active noise control in various noisy environments. *Engineering Applications of Artificial Intelligence* **121**, 105971 (2023).

[18] Mostafavi, A. & Cha, Y.-J. Deep learning-based active noise control on construction sites. *Automation in Construction* **151**, 104885 (2023).

[19] Aboutiman, A., Maamoun, K. S. A., Karimi, H. R. & Ripamonti, F. Hybrid deep learning–based active noise control for encapsulated structures with openings. *Expert Systems with Applications* 131247 (2026). URL https://www.sciencedirect.com/science/article/pii/S0003682X25002956.

[20] Cao, Z. *et al.* Continuous improvement of self-driving cars using dynamic confidence-aware reinforcement learning. *Nature Machine Intelligence* **5**, 145–158 (2023).

[21] Feng, S. *et al.* Dense reinforcement learning for safety validation of autonomous vehicles. *Nature* **615**, 620–627 (2023).

[22] Luo, Z., Ma, H., Shi, D. & Gan, W.-S. Gfanc-rl: Reinforcement learning-based generative fixed-filter active noise control. *Neural Networks* **180**, 106687 (2024).

[23] Ryu, S., Lim, J. & Lee, Y.-S. Narrowband active noise control with ddpg based on reinforcement learning. *International Journal of Automotive Technology* **25**, 1389–1397 (2024).

[24] Raeisy, B. & Haghighi, S. G. Active noise controller with reinforcement learning (2012). Proceedings of the 16th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP 2012).

[25] Sutton, R. S., Barto, A. G. *et al. Reinforcement learning: An introduction* Vol. 1 (MIT press Cambridge, 1998).

[26] Balazs, P., Laback, B. & Majdak, P. *Multiple exponential sweep method for fast measurement of head related transfer functions* (Audio Engineering Society, 2007).

[27] Fujimoto, S., van Hoof, H. & Meger, D. Addressing function approximation error in actor-critic methods. *CoRR* **abs/1802.09477** (2018). URL http://arxiv.org/abs/1802.09477.

[28] Lillicrap, T. P. *et al.* Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971* (2015).

[29] Thrun, S. & Schwartz, A. Issues in using function approximation for reinforcement learning (1993). In: Proceedings of the Connectionist Models Summer School, pp. 255–263.

[30] Hasselt, H. Double q-learning. *Advances in neural information processing systems* **23** (2010).

[31] Konda, V. R. & Tsitsiklis, J. N. Onactor-critic algorithms. *SIAM journal on Control and Optimization* **42**, 1143–1166 (2003).

[32] Mnih, V. *et al.* Asynchronous methods for deep reinforcement learning (2016). In: International Conference on Machine Learning (ICML), pp. 1928–1937.

List of publications

1. **Alkahf Aboutiman**, Reza Shams, Hamid Reza Karimi, Francesco Ripamonti and Marek Pawełczyk. "Active noise control in encapsulated structures with non-minimum phase characteristics using a Kalman filter approach" *Journal of Sound and Vibration*, 615 (2025) 119187.
https://doi.org/10.1016/j.jsv.2025.119187

2. **Alkahf Aboutiman**, Zulfi Rachman, Tin Oberman, Francesco Alletta, Jian Kang, Hamid Reza Karimi and Francesco Ripamonti. "Subjective perception analysis of active noise control algorithms in an encapsulated structure: An experimental study" *Applied Acoustics*, 239 (2025) 110823.
https://doi.org/10.1016/j.apacoust.2025.110823

3. **Alkahf Aboutiman**, Khaled Said Ahmed Maamoun, Hamid Reza Karimi and Francesco Ripamonti. "Hybrid deep learning–based active noise control for encapsulated structures with openings" *Expert Systems With Applications*, 309 (2026) 131247.
https://doi.org/10.1016/j.eswa.2026.131247

4. **Alkahf Aboutiman**, Jan Kirchhof and Francesco Ripamonti. "Transfer Learning-Enhanced DDPG for Adaptive Active Noise Control in Vehicle Interiors" *NeurIPS 2026*, Under Review

5. **Alkahf Aboutiman**, Jan Kirchhof, Hamid Reza Karimi and Francesco Ripamonti. "Deep transfer reinforcement learning for active noise control in vehicle interiors" *Nature Communications*, 2026. Submitted

6. Khaled Said Ahmed Maamoun, **Alkahf Aboutiman**, Ali Mohamed Ali. "Synergistic use of inertial actuators for advanced noise control in transparent patched vibrating plates" *Measurement*, 2025.
https://doi.org/10.1016/j.measurement.2025.118701

7. Chao Liang, **Alkahf Aboutiman**, Jan Kirchhof, Francesco Ripamonti, Hamid Reza Karimi. "Experimental Evaluation of Simultaneous Perturbation Stochastic Ap-

proximation Algorithms for Active Engine Noise Control in Automotive Applications" *Applied Acoustics*.

8. **Alkahf Aboutiman**, Hamid Reza Karimi, Francesco Ripamonti. "A hybrid learning-based active noise control algorithm for encapsulated structures with linear time-varying characteristics" *31st International Congress on Sound and Vibration (ICSV31)*, 2025. Best Paper Award (1st Classified)

9. **Alkahf Aboutiman**, Hamid Reza Karimi, Francesco Ripamonti. "Hybrid Learning-based Active Noise Control in Encapsulated Structures" *International Conference on Machine Learning and Artificial Intelligence for Audio*, 2025

10. **Alkahf Aboutiman**, Hamid Reza Karimi, Francesco Ripamonti. "Frequency-Based Modeling and Active Noise Control Design for Encapsulated Structures" *ICSV30*, 2024