N° d'ordre NNT : $2025ISAL0014$

THESE de DOCTORAT DE L'INSA LYON,
membre de l'Université de Lyon

Ecole Doctorale N° 162
MÉCANIQUE, ÉNERGÉTIQUE, GÉNIE CIVIL, ACOUSTIQUE

Spécialité : MÉCANIQUE

Soutenue publiquement le 17/02/2025, par :

Maryne Febvre

Artificial intelligence to optimize distributed vibration control: Application to transducer networks in smart structures.

Devant le jury composé de :

Berardengo	Marta	Associate Professor	Universita di Genova	Examinatrice
Dion	Jean-Luc	Professeur	Isae Supméca	Rapporteur
Gardonio	Paolo	Professeur	Università degli Studi di Udine	Examinateur
Meyer	Yann	Professeur	Université Savoie Mont Blanc	Président du jury/Rapporteur
Chesne	Simon	Professeur	INSA Lyon	Directeur de thèse
Collet	Manuel	Professeur	Ecole Centrale Lyon	Co-directeur de thèse
Rodriguez	Jonathan	MCF	INSA Lyon	Encadrant de thèse

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Département FEDORA – INSA Lyon - Ecoles Doctorales

SIGLE	ECOLE DOCTORALE	NOM ET COORDONNEES DU RESPONSABLE
ED 206 CHIMIE	**CHIMIE DE LYON** https://www.edchimie-lyon.fr Sec. : Renée EL MELHEM Bât. Blaise PASCAL, 3e étage secretariat@edchimie-lyon.fr	**M. Stéphane DANIELE** C2P2-CPE LYON-UMR 5265 Bâtiment F308, BP 2077 43 Boulevard du 11 novembre 1918 69616 Villeurbanne directeur@edchimie-lyon.fr
ED 341 E2M2	**ÉVOLUTION, ÉCOSYSTÈME, MICROBIOLOGIE, MODÉLISATION** http://e2m2.universite-lyon.fr Sec. : Bénédicte LANZA Bât. Atrium, UCB Lyon 1 Tél : 04.72.44.83.62 secretariat.e2m2@univ-lyon1.fr	**Mme Sandrine CHARLES** Université Claude Bernard Lyon 1 UFR Biosciences Bâtiment Mendel 43, boulevard du 11 Novembre 1918 69622 Villeurbanne CEDEX e2m2.codir@listes.univ-lyon1.fr
ED 205 EDISS	**INTERDISCIPLINAIRE SCIENCES-SANTÉ** http://ediss.universite-lyon.fr Sec. : Bénédicte LANZA Bât. Atrium, UCB Lyon 1 Tél : 04.72.44.83.62 secretariat.ediss@univ-lyon1.fr	**Mme Sylvie RICARD-BLUM** Laboratoire ICBMS - UMR 5246 CNRS - Université Lyon 1 Bâtiment Raulin - 2ème étage Nord 43 Boulevard du 11 novembre 1918 69622 Villeurbanne Cedex Tél : +33(0)4 72 44 82 32 sylvie.ricard-blum@univ-lyon1.fr
ED 34 EDML	**MATÉRIAUX DE LYON** http://ed34.universite-lyon.fr Sec. : Yann DE ORDENANA Tél : 04.72.18.62.44 yann.de-ordenana@ec-lyon.fr	**M. Stéphane BENAYOUN** Ecole Centrale de Lyon Laboratoire LTDS 36 avenue Guy de Collongue 69134 Ecully CEDEX Tél : 04.72.18.64.37 stephane.benayoun@ec-lyon.fr
ED 160 EEA	**ÉLECTRONIQUE, ÉLECTROTECHNIQUE, AUTOMATIQUE** https://edeea.universite-lyon.fr Sec. : Philomène TRECOURT Bâtiment Direction INSA Lyon Tél : 04.72.43.71.70 secretariat.edeea@insa-lyon.fr	**M. Philippe DELACHARTRE** INSA LYON Laboratoire CREATIS Bâtiment Blaise Pascal, 7 avenue Jean Capelle 69621 Villeurbanne CEDEX Tél : 04.72.43.88.63 philippe.delachartre@insa-lyon.fr
ED 512 INFOMATHS	**INFORMATIQUE ET MATHÉMATIQUES** http://edinfomaths.universite-lyon.fr Sec. : Renée EL MELHEM Bât. Blaise PASCAL, 3e étage Tél : 04.72.43.80.46 infomaths@univ-lyon1.fr	**M. Hamamache KHEDDOUCI** Université Claude Bernard Lyon 1 Bât. Nautibus 43, Boulevard du 11 novembre 1918 69 622 Villeurbanne Cedex France Tél : 04.72.44.83.69 direction.infomaths@listes.univ-lyon1.fr
ED 162 MEGA	**MÉCANIQUE, ÉNERGÉTIQUE, GÉNIE CIVIL, ACOUSTIQUE** http://edmega.universite-lyon.fr Sec. : Philomène TRECOURT Tél : 04.72.43.71.70 Bâtiment Direction INSA Lyon mega@insa-lyon.fr	**M. Etienne PARIZET** INSA Lyon Laboratoire LVA Bâtiment St. Exupéry 25 bis av. Jean Capelle 69621 Villeurbanne CEDEX etienne.parizet@insa-lyon.fr
ED 483 ScSo	**ScSo**[1] https://edsciencessociales.universite-lyon.fr Sec. : Mélina FAVETON Tél : 04.78.69.77.79 melina.faveton@univ-lyon2.fr	**M. Bruno MILLY** (INSA : J.Y. TOUSSAINT) Univ. Lyon 2 Campus Berges du Rhône 18, quai Claude Bernard 69365 LYON CEDEX 07 Bureau BEL 319 bruno.milly@univ-lyon2.fr

[1] ScSo : Histoire, Géographie, Aménagement, Urbanisme, Archéologie, Science politique, Sociologie, Anthropologie

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Remerciements

Au terme de ces années de doctorat, il m'est essentiel d'exprimer ma profonde gratitude à celles et ceux qui ont contribué(e)s, de près ou de loin, à l'aboutissement de cette thèse.

Je tiens tout d'abord à remercier mes directeurs de thèse, Manuel Collet et Simon Chesné ainsi que mon encadrant Jonathan Rodriguez, pour leur encadrement, leur patience et leurs précieux conseils tout au long de ces années. Leur expertise et leur rigueur scientifique ont été une source constante d'inspiration et m'ont permis d'avancer avec confiance dans mes travaux.

Je remercie également les membres de mon jury, Marta Berardengo, Jean-Luc Dion, Paolo Gardonio, Yann Meyer, d'avoir accepté d'évaluer ce travail et pour leurs remarques enrichissantes qui m'ont permis d'en affiner les conclusions. Un immense merci aux rapporteurs, Jean-Luc Dion et Yann Meyer pour le temps consacré à la lecture de cette thèse et pour leurs retours constructifs.

J'exprime ma reconnaissance aux membres de mes deux laboratoires le LaMCoS et le LTDS, et plus particulièrement à mes collègues et ami(e)s doctorant(e)s pour les discussions stimulantes, les échanges scientifiques, les moments de convivialité et les conseils botaniques qui ont rendu ces années plus agréables. Je remercie également les équipes administratives et techniques, pour leurs aides précieuses dans la gestion des aspects pratiques de la recherche.

Sur un plan plus personnel, je tiens à remercier ma famille et mes ami(e)s, qui m'ont soutenu(e) dans les moments de doute et m'ont encouragé(e) jusqu'au bout. Leur présence et leur patience ont été essentielles. Une pensée particulière à celui qui partage ma vie, dont la patience, le soutien et la bienveillance ont été un refuge précieux tout au long de cette aventure.

Enfin, merci à toutes celles et ceux qui, de près ou de loin, ont contribué à cette aventure. Cette thèse est aussi le fruit de votre bienveillance et de votre soutien.

To all those who will take the time to read these pages, thank you for your interest and your curiosity. I hope you will find in this work the spark of the passion that guided me through this journey.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Abstract

The development of metamaterials increasingly incorporates networks of transducers, transforming traditional materials into functional "smart" materials capable of interacting with multiphysical systems. Piezoelectric transducers play a key role in these smart structures, enabling energy harvesting, wave propagation control, and vibration control to address issues such as human discomfort, material fatigue, and instability. These multiphysical elements facilitate the conversion between mechanical and electrical energy, allowing their use as both sensors and actuators. However, the design of such adaptive smart structures becomes more complex with the integration of multiple transducers, requiring the tuning of numerous parameters for optimal performance.

Vibration control in smart structures is traditionally achieved using classical active and passive methods. While effective for systems with a limited number of transducers, these approaches face challenges as system complexity increases. The optimization of control parameters becomes significantly harder, particularly when addressing non-convex problems with physical uncertainties. To overcome these challenges, recent research has focused on intelligent control methods, including artificial intelligence (AI)-based techniques. In line with recent literature, this thesis explores the use of reinforcement learning algorithms using neural networks to adjust control law parameters and mitigate various disturbance signals on experimental smart structures.

In this work, active and passive vibration control using shunted piezoelectric transducers is optimized through Deep Reinforcement Learning (DRL). The performance of DRL is compared with other optimization methods, including analytical approaches, simplex, and genetic algorithms. Efficiency is evaluated in terms of vibration damping, structural stability, and computational effort. Modal analysis, both numerical and experimental, is performed on the various structures to verify their suitability for control using piezoelectric transducers. The DRL methodology is first introduced on a 1D finite element structure. Then experiments are conducted on a smart cantilever beam featuring two collocated piezoelectric transducers. Subsequently, the approach is extended to networks of transducers for distributive control of larger and more complex structures.

The results demonstrate the effectiveness, advantages, and limitations of using AI in vibration control. For DRL to tune effective control laws, components such as observation, action, and reward must be carefully defined in alignment with the control objectives, using metrics representative of the structure's vibration state. DRL successfully tunes multi-parameter feedback control laws, optimizing vibration damping while ensuring stability under varying disturbance signals. Unlike some traditional methods, it allows for non-convex optimization functions. However, due to the randomness inherent in neural network initialization and training, DRL can occasionally lead to sub-optimal solutions or divergence. Sequencing training strategies are being developed that exploit the memory principles of neural networks, thereby increasing the robustness of the method. These strategies also allow the system to adapt to environmental variations, such as changes in disturbances or objective functions.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Resumé

Le développement des métamatériaux intègre de plus en plus des réseaux de transducteurs, transformant les matériaux traditionnels en matériaux "intelligents" fonctionnels capables d'interagir avec des systèmes électromécaniques. Les transducteurs piézoélectriques jouent un rôle clé dans ces structures intelligentes, permettant la récupération d'énergie, la propagation d'ondes et le contrôle des vibrations pour résoudre des problèmes tels que l'inconfort humain, la fatigue des matériaux et l'instabilité des structures contrôlés. Ces éléments multiphysiques facilitent la conversion de l'énergie mécanique et électrique de manière réciproque, ce qui permet leur utilisation à la fois comme capteurs et actionneurs. Cependant, la conception de telles structures intelligentes adaptatives devient de plus en plus complexe avec l'intégration de multiples transducteurs, nécessitant l'ajustement de nombreux paramètres pour atteindre des performances optimales dans une approche souvent décentralisée.

Le contrôle des vibrations dans les structures intelligentes est traditionnellement réalisé à l'aide de méthodes classiques actives ou passives. Bien que ces approches soient efficaces pour des systèmes comportant un nombre limité de transducteurs, elles rencontrent des difficultés lorsque la complexité du système augmente. L'optimisation des paramètres de contrôle devient alors nettement plus difficile, en particulier lorsqu'il s'agit de résoudre des problèmes non convexes avec des incertitudes physiques. Pour relever ces défis, des recherches récentes se sont concentrées sur des méthodes de contrôle intelligent, notamment en utilisant des techniques de contrôle basées sur l'intelligence artificielle (IA). Conformément à la littérature récente, cette thèse explore l'utilisation d'algorithmes d'apprentissage par renforcement basés sur des réseaux de neurones pour ajuster les paramètres des lois de contrôle et atténuer divers signaux de perturbation sur des structures intelligentes expérimentales.

Dans ce travail, le contrôle actif et passif des vibrations à l'aide de transducteurs piézoélectriques shuntés est optimisé grâce à l'apprentissage par renforcement profond (Deep Reinforcement Learning, DRL). Les performances du DRL sont comparées à celles d'autres méthodes d'optimisation, notamment les approches analytiques, le simplex et les algorithmes génétiques. L'efficacité est évaluée en termes d'atténuation des vibrations, de stabilité structurelle et de couts calcul. Des analyses modales, à la fois numériques et expérimentales, sont réalisées sur différentes structures afin de garantir un impact mesurable de l'ajout d'un contrôle à l'aide de transducteurs piézoélectriques. La méthodologie DRL est d'abord introduite sur une structure unidimensionnelle modélisée par éléments finis. Ensuite, des expériences sont menées sur une poutre intelligente encastrée-libre équipée de deux transducteurs piézoélectriques colocalisés. Enfin, l'approche est étendue à des réseaux de transducteurs pour un contrôle distribué de structures plus grandes et beaucoup plus complexes.

Les résultats démontrent l'efficacité, les avantages et les limites de l'utilisation de l'IA pour le contrôle des vibrations. Pour que le DRL puisse ajuster efficacement les lois de contrôle, des composants tels que l'observation, l'action et la récompense doivent être soigneusement définis en accord avec les objectifs de contrôle, en utilisant des métriques représentatives de l'état vibratoire de la structure. Le DRL ajuste avec succès des lois de contrôle en boucle fermée multi-paramètres, optimisant l'atténuation des vibrations tout en garantissant la stabilité face à des signaux de perturbation variables. Contrairement à certaines méthodes traditionnelles, il permet de traiter des fonctions d'optimisation non convexes. Cependant, en raison du caractère aléatoire de l'initialisation et de l'entraînement

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

des réseaux de neurones, le DRL peut parfois conduire à des solutions sous-optimales ou même diverger. Des stratégies d'entraînement séquentiel sont développées pour exploiter les principes de mémoire des réseaux neuronaux, augmentant ainsi la robustesse de la méthode. Ces stratégies permettent également au système de s'adapter aux variations de l'environnement, telles que les changements de perturbations ou des modifications de la fonction objectif.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Contents

1 Introduction **11**
- 1.1 Context . . . 12
  - 1.1.1 Definition of Smart Structures . . . 12
  - 1.1.2 Definition of transducers through the discovery of the piezoelectric effect . . . 15
  - 1.1.3 Definition of artificial intelligence . . . 18
- 1.2 Thesis Framework . . . 19
  - 1.2.1 Motivations . . . 19
  - 1.2.2 Objectives . . . 19
  - 1.2.3 Approaches . . . 19
  - 1.2.4 Exclusion . . . 19
- 1.3 Organisation . . . 20
  - 1.3.1 Chapter 2: From piezoelectric material to smart structure using artificial intelligence. . . . 20
  - 1.3.2 Chapter 3: Introduction to the method: Deep Reinforcement Learning for passive control on a 1D element model. . . . 20
  - 1.3.3 Chapter 4: Tuning active control laws on a smart cantilever beam: experimental study . . . 20
  - 1.3.4 Chapter 5: Tuning methods for distributed active vibration control of a smart structure : an experimental study . . . 21
  - 1.3.5 Chapter 6: Tuning distributed shunt control of a smart structure: numerical study . . . 21
  - 1.3.6 Chapter 7: Conclusions . . . 21

2 From piezoelectric material to smart structure using artificial intelligence. **23**
- 2.1 State of the Art and Equations for Smart Structures . . . 24
  - 2.1.1 Structure Dynamics and Piezoelectricity . . . 24
  - 2.1.2 Control of Piezoelectric Structures . . . 33
  - 2.1.3 Optimization Algorithm . . . 51
- 2.2 Theory of Machine Learning in AI . . . 58
  - 2.2.1 From Perceptron to Neural Network . . . 58
  - 2.2.2 Reinforcement learning . . . 64
- 2.3 Piezoelectric structures and Artificial Intelligence . . . 66
  - 2.3.1 Artificial Neural Network Controller . . . 66
  - 2.3.2 Neural Networks as Controller Optimizer on smart structures . . . 68
- 2.4 Conclusion: Position the subject within the state-of-the-art . . . 69

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

3 Introduction to the method: Deep Reinforcement Learning for passive control on a 1D element model. 71
3.1 Smart cantilever beam model . . . 72
3.1.1 Finite element model equations . . . 72
3.1.2 System response and disturbance . . . 74
3.1.3 Passive Control implementation . . . 75
3.2 Proposed Methodology applied on the example case . . . 76
3.2.1 Environment Definition . . . 77
3.2.2 Reinforcement Learning Algorithm . . . 79
3.3 Control Performances . . . 80
3.4 Conclusion . . . 81

4 Tuning active control laws on a smart cantilever beam: experimental study 83
4.1 Definition and design of a smart cantilever beam with two collocated piezoelectric patches . . . 84
4.1.1 Finite element model . . . 84
4.1.2 Experimental validation . . . 88
4.2 Active control definition . . . 94
4.2.1 Applying the control to the structure . . . 94
4.2.2 Parametric tuning . . . 95
4.2.3 Disturbance signals . . . 96
4.3 Deep reinforcement learning applied to the system . . . 96
4.3.1 Environment definition . . . 96
4.3.2 Reward . . . 98
4.3.3 Algorithm . . . 100
4.4 Results and performances . . . 104
4.4.1 Derivative controller . . . 104
4.4.2 Lead Controller . . . 107
4.4.3 Damping effect on the test bench . . . 111
4.5 Discussion . . . 116
4.5.1 Benefits and limits of DRL approach . . . 116
4.5.2 Stability concern . . . 117
4.5.3 Algorithm robustness . . . 117
4.6 Conclusion . . . 122

5 Tuning methods for distributed active vibration control of a smart structure : an experimental study 123
5.1 Definition of a smart multi-input multi-output (MIMO) system . . . 125
5.1.1 Design of a smart beam . . . 125
5.1.2 Experimental implementation . . . 128
5.1.3 Implementing control: modeling and experimental comparison . . . 133
5.2 Control tuning methods . . . 136
5.2.1 Environment definition . . . 136
5.2.2 Methods comparison . . . 139
5.2.3 Control performances on the model . . . 143
5.3 Robustness of tuning methods to reward variation . . . 145
5.3.1 Target frequency bandwidth expansion . . . 145
5.3.2 Stability criteria improvement . . . 149

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

5.3.3 Control performances on the test bench . . . 152
5.4 Conclusion . . . 154

6 Tuning distributed shunt control of a smart structure: numerical study 155
6.1 Definition of a smart multi-input multi-output system . . . 157
6.1.1 Introduction to the study structure . . . 157
6.1.2 Model definition of the smart beam . . . 159
6.1.3 Implementing control . . . 166
6.2 Control tuning methods . . . 167
6.2.1 Environment definition . . . 167
6.2.2 Methods comparison . . . 170
6.2.3 Optimisation results on the modelled beam . . . 174
6.3 Discussion and Conclusion . . . 183

7 Conclusions 185
7.1 Main findings . . . 186
7.2 Main results . . . 186
7.3 Implementing deep reinforcement learning for control optimization: Main steps . . . 188
7.3.1 Preliminary questions . . . 188
7.3.2 Selection of the AI algorithm . . . 189
7.3.3 Element Definitions . . . 190
7.3.4 Training . . . 192
7.3.5 Results . . . 192
7.4 Ecological Cost . . . 193
7.5 Future works: Going further with mechanical applications . . . 193
7.5.1 Short term: Mechanical tools . . . 193
7.5.2 Mid-term: Use full capacities of AI tools . . . 193
7.5.3 Long-term: Adapt AI tools for mechanical applications . . . 193
7.5.4 Last words . . . 194

A 1D Finite Element Model of a Smart Cantilever Beam: Numerical Datas 207

B Cantilever beam transfer function values for Chapter 4 209

C Measured and Identified Distributed System Transfer Function from Chapter 5 211

D Modal distribution of the structure between 0 and 590 Hz: modal shapes 217

E Shunt RC_{neg} Equivalent capacity demonstration 221

F Calculation of the $C0_2$-equivalent of the numerical tools used 225

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Chapter 1

Introduction

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

This chapter introduces the context of the thesis by defining each term used in the title. Then, the motivations, objectives, and approaches are explained. Finally, an overview of the following chapters is presented.

1.1 Context

In the late 20th century, a new field of research began to emerge involving multiphysical materials applied to structures at different scales. These materials have internal properties that allow them to respond quickly, autonomously, reversibly, and predictably to external stimuli such as temperature, voltage, magnetic fields, humidity, light, pressure, etc. These so-called smart-materials will act as sensor and actuator networks within structures, making them adaptable. Newly created structures come with the need to understand global multiphysics interaction. The goal is to create and use structures with a defined function, such as a mechanical diode, wave trap, or vibration insulator. New methodologies to properly tune the multiphysical material networks are investigated. In this thesis it is chosen to use an Artificial Intelligence (AI) tool to tune transducers within a network in order to program already designed smart structures. The following paragraphs propose a definition of each keyword.

1.1.1 Definition of Smart Structures

Smart structures, also known as intelligent, adaptive, active or sensory structures, are systems designed to integrate actuators, sensors and control units into materials or components to achieve adaptive and multifunctional performance. Originally inspired by biological principles, they mimic the stimulus-response characteristics of living organisms [37]-[144]. Giurgiutiu [37] proposed two paradigms for understanding smart structures: the **technological paradigm**, which focuses on integrating actuators, sensors, and controls within a structural component, and the **science paradigm**, which aims to create biologically inspired materials with intelligence and life-like features embedded at the microstructural level. Anderson [3] emphasized their ability to respond proportionally to external stimuli to mitigate undesired effects or enhance desirable ones, while Spillman [59] defined them as non-biological physical systems with a specific purpose, mechanisms to achieve it, and a biological-like functionality pattern.

Smart structures can be classified based on their complexity and functionality. Hassani [45] outlined a hierarchical classification: **sensory structures**, equipped with sensors for real-time monitoring; **adaptive structures**, incorporating actuators for controlled excitation; **controlled structures**, combining sensors and actuators within a feedback system; **active structures**, which integrate advanced sensing and actuation; and **intelligent structures**, which employ sophisticated control systems for enhanced damage monitoring and response. Hurlebaus [53] and Spillman [59] highlighted the synonymous use of terms such as intelligent structures, adaptronics, and structronics in describing systems that integrate actuators, sensors, and control units to achieve adaptive functionality.

The capabilities of smart structures extend beyond traditional structural functions. They are designed to perform tasks such as vibration isolation, acoustic noise reduction, precision alignment, shape modification, and structural health monitoring, all while sustaining mechanical loads [144]-[45]. Additionally, they can achieve sensing, actuation, autonomous power generation, and signal processing without significantly increasing mass

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

or power consumption [53]-[37]. These multifunctional characteristics make smart structures particularly valuable in advanced applications, such as high-aspect-ratio adaptive wings for aerospace systems [3].

According with this definition, one of the oldest and most famous adaptive and programmable mechanisms in aerospace was imagined by Leonardo Da Vinci and drawn in the Cortex Anticus in 1496. The mechanism aims to reproduce the bidding wing behavior through different flight stages (Fig.1.1a).

The work of fluid mechanics during the 19th and 20th centuries allows to conceive aerodynamic wings geometry to adapt to the longest flight phase. A new research area studies deformable and adaptable wings to reduce energy losses during all flight phases (Fig.1.1c). These wings are instrumented with sensors and actuators such as motors that allow shape variation in coherence with the environment (temperature, pressure, speed, etc.).

From the point of view of sustainable development, new airplane technologies are emerging that do not require fuel, such as Solar Impulse (Fig.1.1b). The mass and geometry of the aircraft are designed for optimal aerodynamics. Energy consumption is reduced through the use of photovoltaic cells. Considering the already optimized internal shape, the use of an adaptive system for vibration control offers a new perspective to improve the energetic balance, adapted to variable flying conditions. Mass addition and energy balance remain critical challenges to consider when seeking performance.

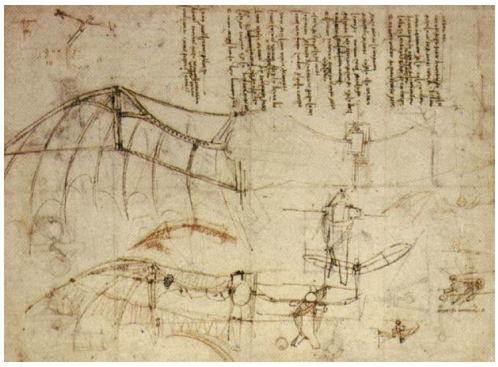
At the beginning of the 21st century, the use of " smart " arrangements of material in space was introduced. These shapes are integrated into structures such as airfoils to offer new structural morphing concepts such as improved deformation characteristics. Figure 1.1d shows examples of truss cores designed and studied to support large deflections without exceeding the yield strength.

No longer just bio-inspired, smart structures are being used at various scales and are finding significant applications in devices such as telescopes equipped with metalenses (Fig. 1.2a and Fig. 1.2c) or piezoelectric metamirrors [120]. These components can be designed to be deformable and adaptable (Fig. 1.2b). With adaptive materials, sensors and actuators do not need to be added via a mechanism, but can be integrated directly into structures.

Current research efforts are focused on developing new generations of adaptive structures and systems composed of distributed composites that integrate functional materials such as energy focusing with mechanical lens (Figure 1.3a), mechanical diodes (Figure 1.3b) or mechanical insulator (Figure 1.3c and 1.3d). These efforts aim to establish a suite of methodological tools, numerical approaches, and technologies that facilitate a paradigm shift in the classical problem of active control in vibratory systems.

Instead of relying on extensive theoretical frameworks to achieve the "optimal" control of distributed structures — which are often inherently uncontrollable — some approaches focus on optimizing simple control laws. These laws are applied to dense networks of distributed transducers, enabling the emulation of effective, quasi-colocated local control strategies. Aligning with this perspective, this thesis proposes new strategies for optimizing control laws specifically tailored to distributed transducers.

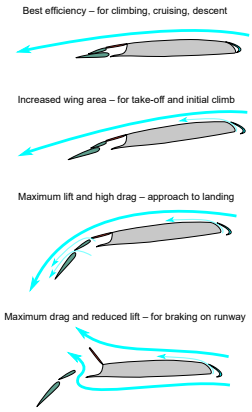
Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés



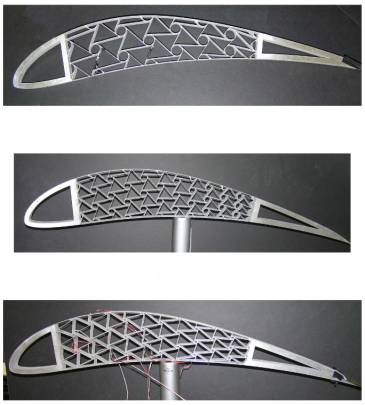
(a) By Leonardo Da Vinci: Drawing of a flying machine made in 1488



(b) By Rama: Solar Impulse is the first fuel-free aircraft to circumnavigate the globe between 2015 and 2016.

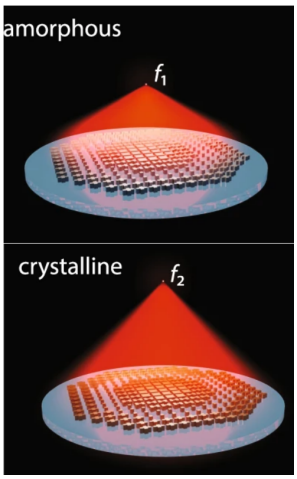


(c) By Andrew Fry: Variation in aircraft wing geometry during different phases of flight

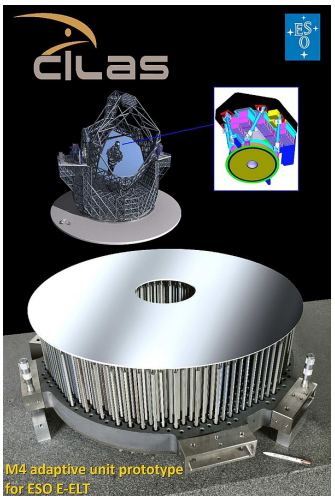


(d) By Spadoni [133]: Manufactured truss-core airfoils (Numerical and experimental analysis in 2007).

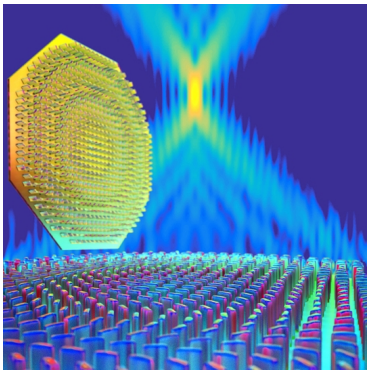
Figure 1.1: Aircraft wing design evolution with time through four examples



(a) By Mikhail Y. Shalaginov and al. [124]: Artistic rendering of a reconfigurable varifocal metalens



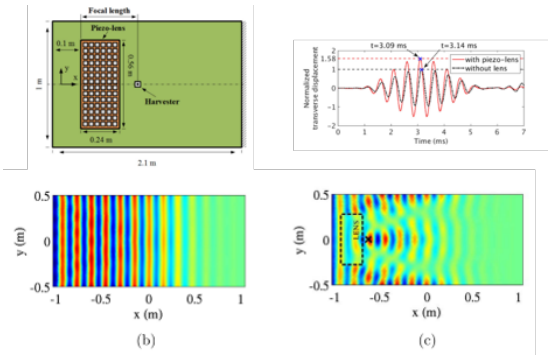
(b) By Cilas [22]: Prototype of the M4 deformable mirror for E-ELT



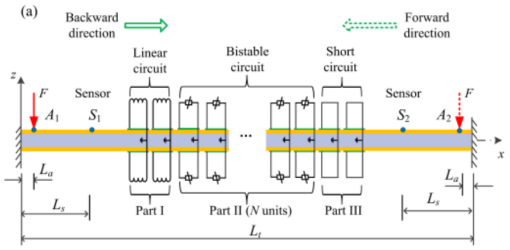
(c) By Mohammadreza Khorasaninejad and al [63]: An illustration of the ultra-thin planar lens. focussing an incident light to a spot smaller than the wavelength. (Image courtesy of Peter Allen)

Figure 1.2: Example of adaptive structures applied to optical technologies.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés



(a) From K. Yi et al [151] : Piezoelectric lens



(b) From: Zheng Yisheng et al [153]: Piezoelectric smart-structure for adaptive nonreciprocal wave transmission

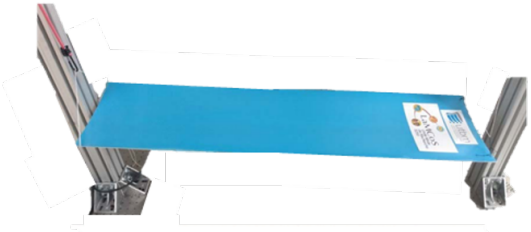
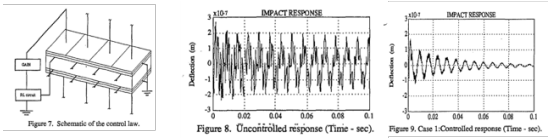


Fig. 4 Smart composite structure: global view

(c) From: J.Rodriguez et al [116]: Piezoelectric smart-structure with active vibration control



(d) From: H. Cudney et al [23]: Multilayered piezoelectric actuators for distributive control

Figure 1.3: Piezoelectric smart structures literature overview

1.1.2 Definition of transducers through the discovery of the piezoelectric effect

Piezoelectricity is found in many everyday objects, such as lighters, microphones, or measuring tools. The etymology of the word piezoelectric uses a prefix from the Greek πιέζειν (piézein), which means to squeeze or press. The term electricity defines a form of energy characterized by the movement of charges. Thus, piezoelectricity is the generation of electricity by applying pressure to a material that results in small deformation.

In 1880, the discovery of piezoelectric material is made by Pierre and Jacques Curie [24](p.6). They demonstrate experimentally that non-conducting crystals become electrical dipoles under compressive mechanical stress. They use Thomson's electrometer to measure the electrical potential difference and observe a change in polarity as the material is stretched or compressed named as the direct piezoelectric effect. Lippman demonstrates indirect piezoelectricity analytically using the conservation law of electricity in 1881 [79]. He assumes that the crystal can stretch proportionally to an applied voltage, and the experimental proof is demonstrated in the same year by the Curie brothers [24](p.26 et p.40). The experimental setup used consists of two superposed piezoelectric subsystems constrained between two fixed plates (Figure 1.4). The first system (a, b, c) measures pressure variations by being connected to an electrometer (e). The second system (a', b', c'), which produces the phenomenon under study, is connected to a Holtz machine (M). This electrostatic high-voltage generator uses a fast-rotating disk in contact with a fixed disk through a conductive comb. With this setup, it is possible to observe the contraction and expansion cycles of the piezoelectric crystals. The relationship between the direction of the applied voltage and the dilation and contraction of the material is thus identified.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

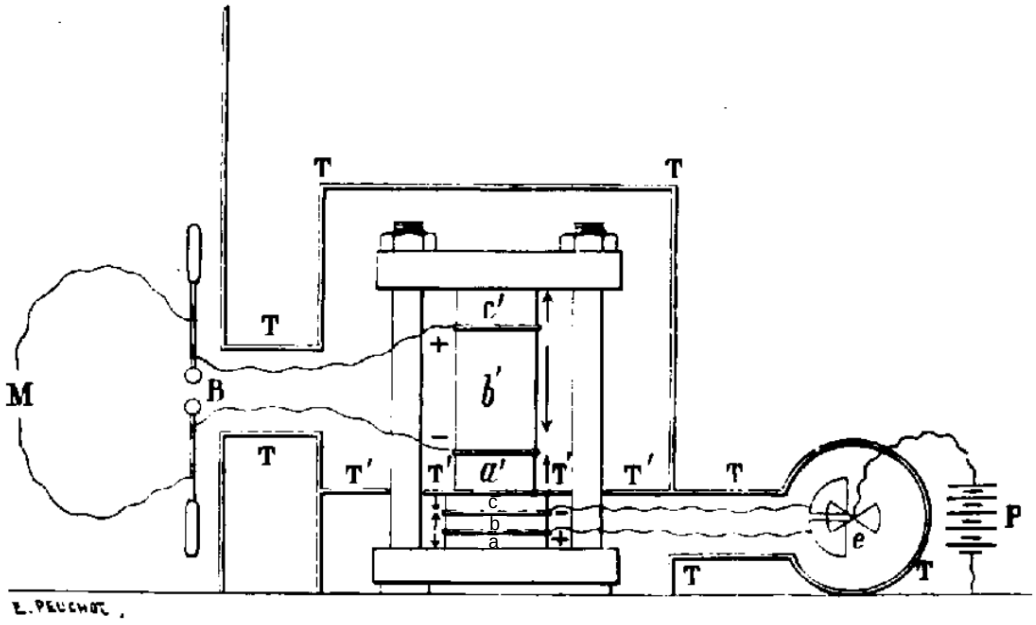


Figure 1.4: Diagram of the experimental set-up used by the Curie brothers to demonstrate the indirect piezoelectric effect [24] (p.40): pressure measurement system (a, b ,c), system producing the phenomenon (a', b', c'), Holtz machine (M), ball micrometer poles (B), ground (T and T'), battery (P) of the eletrometer (e).

Piezoelectric material fits perfectly within the definition of transducers since it can transfer energy from the mechanical to the electrical domain with a reciprocity effect and they can be considered as smart materials.

Smart materials are advanced material systems that exhibit the ability to sense, respond, or adapt to changes in their environment in a manner analogous to biological systems. Giurgiutiu [37] described them as "evolutionary material systems" with sensing and actuation properties similar to those observed in the living world. Unlike smart structures, where sensors, actuators, and control mechanisms are externally integrated, smart materials incorporate these elements directly into their microstructure, enabling inherent functionality [3]. According to Vepa [144], smart materials possess properties that can be significantly influenced or modified by environmental conditions such as temperature, pressure, or electromagnetic fields. A classification depending on the coupling behavior of different multiphysical materials is proposed in Figure 1.5 inspired by the work of Thangudu [138]. As noted by Hurlebaus [53], the term "smart" does not imply an intrinsic intelligence within the materials but rather the exploitation of their unique properties to achieve superior design functionality compared to conventional materials.

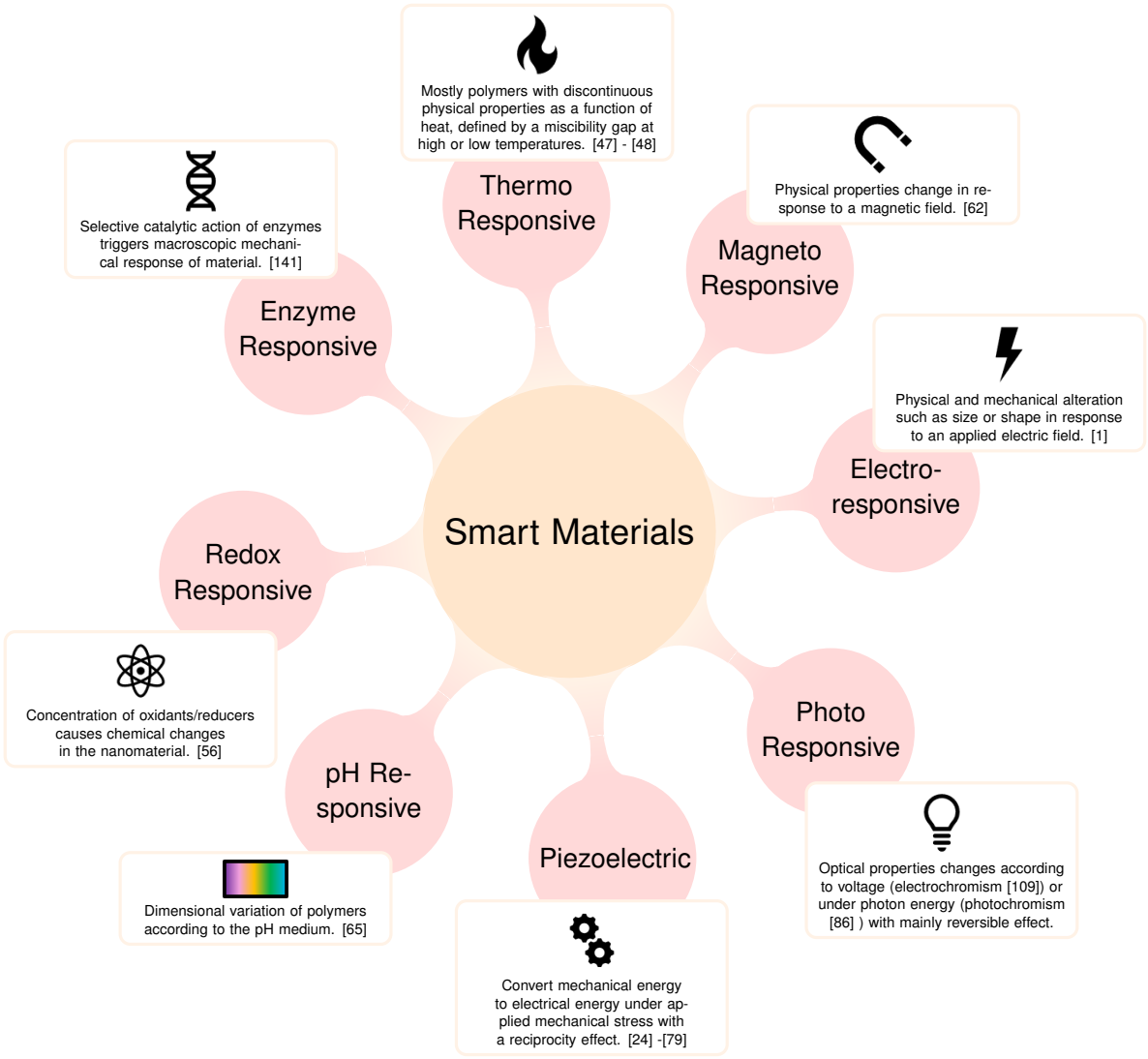


Figure 1.5: Classification of smart materials based on the physics of external stimuli

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

1.1.3 Definition of artificial intelligence

From a definition point of view, the first meaning of the word "intelligence" is about the human cognitive ability to understand, adapt, and think. The word "artificial" has two meanings. The first refers to human production as opposed to nature. The second meaning refers to something that does not correspond to reality.

The French National Center for Textual and Lexical Resources (CNRTL) proposes a definition of "artificial intelligence" as follows: research into ways of making computers comparable to human intellectual abilities. This definition is not complete since intellectual ability here refers to results, not the means of achieving them. An algorithm can analyze a considerable amount of information in a relatively short amount of time compared to a human, but the human needs fewer samples to build his knowledge.
This definition can also refer to the imitation of the human brain by using neural networks. For example, Alan Turing questioned the consciousness of robots in 1950 with a human conversation imitation test [140]. However, this method is not always used in all AI algorithms. The advancement of algorithms is not anymore aimed at imitating humans. The technology is used as an analytic tool or to perform repetitive tasks. Reliability of results, speed, and execution of complex tasks are critical factors in algorithm design. These properties of AI algorithms must be studied in the context of distributed smart structures.

Today, the tools of artificial intelligence are bringing very powerful advantages to daily activities. International organizations (UNESCO, European Union (EU)) promote the ethics and responsible use of AI. It is more globally involved in the computer and robotics ethics, initiated by the three laws of Asimov [4]:

- A robot cannot injure a human being, even by being passive, thus exposing the human being to danger.
- A robot must obey the orders of a human being, except when it conflicts with Law Number One.
- A robot must protect its existence if it does not contradict the first two laws.

Thus, the robot's duties are not to harm humans, to obey, and to protect its own life. More recently, the protection of the environment has been registered into the ethics with mandatory energy sobriety.

When it comes to AI, ethics covers a broader area with more technical aspects. For example, the bias of the algorithms coming from the databases used for training must be considered. Automated decision-making, privacy, and accountability must be defined for each application of the technology. In addition, the energy consumption of AI is not negligible during the three main phases of its life: training, storage, and use. This aspect will be discussed in the conclusion of this thesis.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

1.2 Thesis Framework

1.2.1 Motivations

This manuscript specifically develops the use of piezoelectric materials as transducers to control vibrations of thin structures. Organized in a network, their use introduces an experimental limit to master multiphysical behaviors, especially when adding distributive control laws. Uncertainties due to the smart structure can occur, leading to instability or inefficient vibration control.

1.2.2 Objectives

This thesis chooses to investigate optimization tools developed in the field of Artificial Intelligence to tune robust and efficient control laws for vibration mitigation. Energy distribution is controlled by the active or passive use of piezoelectric transducers to create functional structures.

1.2.3 Approaches

This work utilizes reinforcement learning (RL) algorithms to optimize multiple control law parameters in experimental setups. The chosen strategy introduces increasing complexity by incorporating multiple smart structures and progressively expanding both the size of the transducer network and the complexity of the control laws. The progression begins with a structure featuring a single transducer and advances to configurations with 6 and 40 transducers, accompanied by a growing number of control parameters to optimize.

Classical optimization methods, such as parametric optimization, simplex algorithms, and genetic algorithms, are employed as benchmarks to evaluate and compare the tuning performance of the RL-based approach. To facilitate the training of these algorithms, numerical models are developed using finite element modeling or transfer function identification techniques.

1.2.4 Exclusion

This thesis presents a novel approach by bringing together diverse research topics with a primary focus on the study of structural dynamics. Artificial intelligence and automated strategies are used as tools to address vibration control challenges. The aim is not to develop new AI algorithms or invent new control laws, but to explore innovative strategies for applying AI in the context of structural vibration. This research is for academic purposes only and has no commercial intent.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

1.3 Organisation

The thesis is organized as follows:

1.3.1 Chapter 2: From piezoelectric material to smart structure using artificial intelligence.

This chapter presents the interconnection between different research fields, including structural dynamics, piezoelectricity, control systems, and artificial intelligence. It begins by developing the piezoelectric equations that link mechanical and electrical equations. A literature review follows, highlighting the main active and passive control laws, as well as different tuning methods, with a particular focus on those employed in this thesis. The chapter also covers optimization strategies, detailing the use of gradient descent, line search, simplex, and genetic algorithms. A brief history of artificial intelligence is provided, along with an overview of the development of neural network constitutive equations and machine learning algorithms. The reinforcement learning algorithms (Trust Region Policy Optimization - TRPO and Proximal policy optimization - PPO) utilized in this thesis are described in detail. Finally, several examples of AI applications in smart structures are presented and discussed, positioning the thesis within the literature.

1.3.2 Chapter 3: Introduction to the method: Deep Reinforcement Learning for passive control on a 1D element model.

This chapter presents the proposed approach using Reinforcement Learning (RL) to tune controller parameters. A 1-D smart cantilever finite element model with two piezoelectric transducers is used to illustrate the process. Both passive and active control strategies are implemented to mitigate vibrations. The passive approach involves the use of an electrical shunt with a tunable resistive shunt between the transducer electrodes. The active approach focuses on tuning a derivative controller. The RL-based tuning for both control strategies is compared with the analytically defined optimal tuning, as presented in Chapter 2, to highlight its efficiency.

1.3.3 Chapter 4: Tuning active control laws on a smart cantilever beam: experimental study

This chapter outlines the implementation of a reinforcement learning algorithm for tuning a controller on a cantilever beam instrumented with two piezoelectric patches. Initially, a numerical finite element model is developed to design the experimental structure and optimize the piezoelectric coupling coefficient through modal analysis. An active control method is then applied to reduce vibrations, utilizing active feedback control laws between the two collocated piezoelectric patches. Various optimization methods are employed to tune the parameters of these control laws, including a parametric approach, simplex method, and reinforcement learning. The implementation of the reinforcement learning method within this context is described in detail. Finally, a comparison of the different approaches is presented, focusing on their robustness to external disturbances modifications.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

1.3.4 Chapter 5: Tuning methods for distributed active vibration control of a smart structure : an experimental study

This chapter implements the reinforcement learning (RL) optimization method for active control on a structure composed of transducers arranged in networks. The increase in dimension aims to test the effectiveness of RL for distributed control in single input, single output (SISO) loops. Active control is applied to a free beam with six piezoelectric patches. A numerical modal analysis is performed to define the control objectives for a specific mode and to design the test bench. An experimental model with a multi-input, multi-output (MIMO) state space configuration is identified from the test bench and used to train the optimization algorithms. The positive position feedback control law is tuned using RL and simplex optimisation. Their experimental efficiencies on the structure are tested and compared to attest the efficiency of the RL method. The robustness of the RL method is assessed by varying the optimization objectives, such as increasing the optimization frequency range or imposing stability constraints.

1.3.5 Chapter 6: Tuning distributed shunt control of a smart structure: numerical study

This chapter implements the proposed methods using shunt control on a large-dimensional beam composed of transducers arranged in networks. A free-free beam, instrumented with 20 pairs of patches, is controlled using numerically tunable shunts, including resistance and negative capacitance. Experimental measurements are conducted to assess the accuracy of the finite element model of the beam, which is used to train the reinforcement learning algorithm. The training is divided into step with increasing tuning complexity starting with one parameter to tune, where all transducers are set with the same value of R shunt to 40, where all the RC_{neg} shunt values are different.

1.3.6 Chapter 7: Conclusions

This chapter summarizes the key findings of the research and suggests potential directions for future work. In addition, it provides a reflection on the calculation costs incurred during this study.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Chapter 2

From piezoelectric material to smart structure using artificial intelligence.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

2.1 State of the Art and Equations for Smart Structures

2.1.1 Structure Dynamics and Piezoelectricity

Piezoelectric materials can be used as sensors (acceleration, pressure, force, torque, ...) or actuators (piezoelectric stack actuators, amplified piezoelectric actuators (APA), ...) because of their physical properties [119]. If well designed, they can also be used for energy harvesting [146] with low electrical power at the scale of nanowatts to milliwatts. This energy can be reused in low power electrical devices such as lights or wireless sensors in the perspective of the Internet of Things (IOT).

2.1.1.1 Piezoelectric Effect

Piezoelectric are multiphysical materials able to transfer energy from the mechanical to the electrical domain with a reciprocity effect, the so-called direct and converse piezoelectric effect (Fig. 2.1). The material has non-conducting surfaces (Fig. 2.1-a).
The direct piezoelectric effect (Fig. 2.1-b) [24] is defined by the ability of certain crystals to generate a quantity of electrical charge proportional to an applied external force. In the case of compression (Fig. 2.1-b), the barycenters of positive and negative electrical charges are no longer collocated, leading to a change in material polarization. Changes in internal polarisation induce charge movement in the circuit connected to the electrodes while the material is under stress. The displacement of free surface charges induces a current.
The reverse piezoelectric effect (Fig. 2.1-c) [79] occurs when the material is subjected to an electric field that changes the polarization of the surface charges. The material expands or contracts due to the crystal deformation for charge equilibrium. In the absence of external forces, extension occurs when the polarity between the crystal and the applied voltage is the same (Fig. 2.1-c). If the polarity is different, the material will contract.

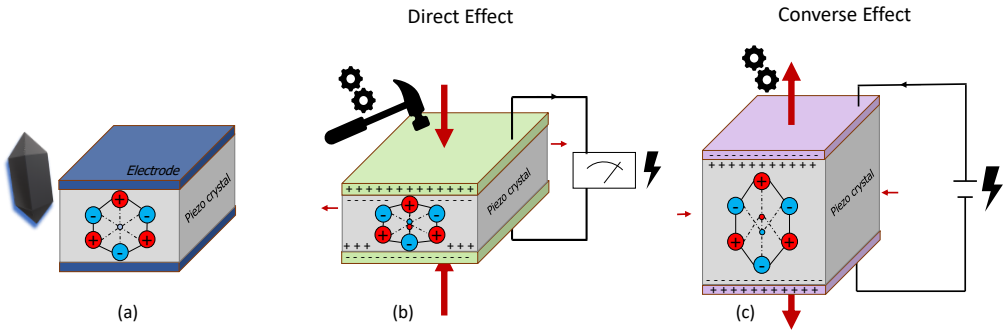


Figure 2.1: Piezoelectric material with no strain (a), shrinkage under stress with direct effect (b), and strain under applied voltage with converse effect (c).

The piezoelectric effect can vary with the direction considered, so it is anisotropic material and occurs in crystal materials with no inversion symmetry. Five types of piezoelectric materials exist [84]: natural crystals (quartz, Rochelle salt, ...), synthetic ceramics ($LiNbO_3$, lead zirconate titanate (PZT), ...), lead-free ceramics ($BiFeO_3$, $Bi_4Ti_3O_{12}$,

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

...), bio-materials (tendon, DNA, ...) and polymers (polyvinylidene fluoride (PVDF), PLLA, ...).

Each piezoelectric material has a defined Curie temperature. At this temperature, the crystal, which is composed of randomly oriented electric dipoles, becomes a zero charge dipole at a macroscopic scale. During the polarization process [103], the crystal is heated and then cooled while exposed to a high electric field. The alignment of microscopic dipoles results on a macroscopic electric dipole. After polarization, the dipoles remain aligned and the material becomes permanently piezoelectric at ambient temperature. A material can lose its piezoelectric properties in two ways: when exposed to a temperature above its Curie temperature, or when exposed to a high electric field of opposite polarization to its own.

2.1.1.2 Equations of a Multi-Physics System

The physics used to describe the phenomenon of piezoelectric systems comes from two domains: mechanics and electrostatics. The combination of these two gives a new writing formalism with the piezoelectric differential equations given by the IEEE Standard [92] in Table 2.1.

Table 2.1: Constitutive Piezoelectric Equations from the IEEE Standard [92]

Mechanical		Electrical	
Stress equation of motion:	$T_{ij,i} = \rho \ddot{u}_j$ (2.7)	Electric field:	$E_i = -\phi_{,i}$ (2.13)
Strain tensor:	$S_{ij} = \frac{1}{2}(u_{i,j} + u_{j,i})$ (2.9)	Electrostatic equation for an insulator:	$D_{i,i} = 0$ (2.20)
Piezoelectric constitutive equations:			
Actuation:	$T_{ij} = c^E_{ijkl} S_{kl} - e_{kij} E_k$ (2.27)		
Sensing:	$D_i = e_{ikl} S_{kl} + \epsilon^S_{ik} E_k$ (2.28)		

Partial differentiation with respect to a space coordinate is indicated by a comma followed by an index, $ui, j = \frac{\partial ui}{\partial xj}$. Variables definitions are described in Table 2.2.

Table 2.2: Piezoelectric equation variables definition

Variable	Name	Unit
Stress Tensor	T_{ij}	$[N.m^{-3}]$
Density	ρ	$[kg.m^{-3}]$
Infinitesimal mechanical displacement	u_j	$[m]$
Strain Tensor	S_{ij}	$[-]$
Electric field intensity	E_i	$[V.m^{-1}]$
Scalar electric potential	ϕ	$[V]$
Electric displacement	D_i	$[C.m^{-2}]$
Elastic constant under $E = 0$	c_{ijkl}	$[N.m^{-2}]$
Piezoelectric constant	e_{kij}	$[C.m^{-2}]$
Relative permittivity under constant Strain	ϵ^S_{ij}	$[-]$

Mathematical demonstrations are given in Tiersen's book [139], but some key reminders and hypotheses are developed in the following lines.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Stress equation of motion:
Consider a moving body as shown in Figure 2.2 to formulate general mechanical equations of motion with their corresponding hypotheses:

Hypothesis 1 *The following hypotheses are used to obtain the stress equations of motion:*

1. *Two types of forces are considered: body forces* $\boldsymbol{f}$ *and contact forces* $\boldsymbol{t}$.
2. *Electrical and magnetic currents acting on the body are negligible.*
3. *Body and surface torques due to gravity are negligible.*
4. *A non-zero volume must be considered*
5. *Small deformations and disturbances are considered in the case of linear application.*

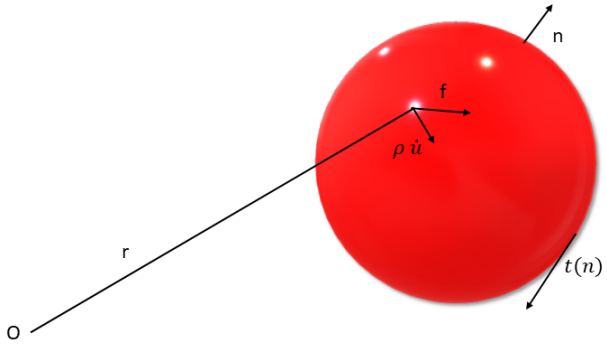


Figure 2.2: A body moving in reference space O, where r is the distance from the origin of reference O, ρ is the density of the body, $\dot{u}$ is the velocity vector, f is the body force vector, n is the normal vector, and $t(n)$ is the contact force vector.

The neglected quantities are relatively small, considering the linear theory. Body forces **f** result from long-range effects such as gravity, electric or magnetic currents and are defined per unit volume. Conversely, contact forces $\mathbf{t(n)}$ result from traction and compression forces on the surface of nearby elements at the macroscopic scale. On the microscopic scale, it is the sum of the interactions of the neighboring elements that lead to small deformations on a global scale.

The conservation equations describing the motion of the homogeneous assembly are as follows

- Conservation of mass over time:

$$(\frac{d}{dt}) \int_V \rho dV = 0 \tag{2.1}$$

- Conservation of linear momentum:

$$\int_S t(n) dS + \int_V f dV = (d/dt) \int_V \rho \dot{u} dV \tag{2.2}$$

Where ρ is the density in $[kg.m^{-3}]$, V is the volume in $[m^3]$, $\dot{u}$ is the speed of a point in $[m.s^{-1}]$, S is the area surrounding the volume in $[m^2]$, $\mathbf{t(n)}$ is the traction-compression vector in $[N.m^{-2}]$, **f** are body forces in $[N.m^{-3}]$, r is the distance between the reference point O and the point on the body in [m].

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Using the divergence phenomenon, the strain tensor $t(n) = n \cdot \overline{\overline{T}}$ can be inserted into the conservation of linear momentum equation, such as:

$$\int_V \overrightarrow{\nabla} \cdot \overline{\overline{T}} dV + \int_V f dV = \int_V \rho(dv/dt) dV \tag{2.3}$$

According to the hypothesis 1.4, the equation (2.3) can be written as:

$$\nabla \cdot T + f = \rho \frac{dv}{dt} \tag{2.4}$$

The first-order linear approximation is used with the hypothesis 1.5, such as:

$$\frac{dv_j}{dt} = \frac{\partial v_j}{\partial t} = \frac{\partial^2 u_j}{\partial^2 t} = \ddot{u}_j \tag{2.5}$$

The equation 2.4, with 3 equations and 9 variables, is written with Cartesian components and considering u as the mechanical displacement, such as:

$$T_{ij,i} + f_j = \rho \ddot{u}_j \tag{2.6}$$

With $f_j = 0$, the stress equation of motion is written such as:

$$\boxed{T_{ij,i} = \rho \ddot{u}_j} \tag{2.7}$$

Strain Tensor :
The system under study is moving, and the material deformation is taken into account with the strain tensor S_{ij} under the hypotheses 2.

Hypothesis 2 *The following hypotheses are used to obtain the strain equations:*

1. *The deformations are continuous.*
2. *The deformation is a bijection of the initial material domain into the deformed material domain.*
3. *Two adjacent elements in the original configuration are adjacent in the deformed configuration.*
4. *Elements inside a closed surface in the initial state remain inside in the deformed state.*
5. *Perturbations are assumed to be small, so displacements are negligible compared to the transformation gradient.*
6. *Small deformations are assumed to be linear, so u^2 terms are neglected.*
7. *Initial and deformed configurations are almost identical ($\frac{\partial A}{\partial \underline{X}} = \frac{\partial A}{\partial \underline{x}}$).*

On Figure 2.3, the displacement of a material point M at a time t is determined by a vector $\overline{U}$ such as:

$$\overline{U} = \overrightarrow{M_0 M_t} = \overrightarrow{M_0 0} - \overrightarrow{0 M_t} = \overline{x} - \overline{X} \tag{2.8}$$

By definition, the displacement gradient tensor is expressed as : $u_{i,j} = \frac{\partial U_i}{\partial X_j}$.
The second order displacement tensor $u_{i,j}$ can be decomposed into the sum of a symmetric tensor S_{ij} and an antisymmetric tensor ω_{ij} such as:

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

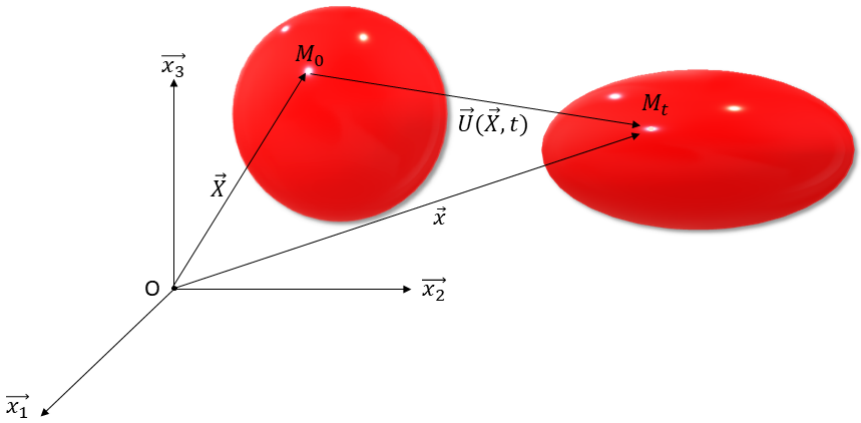


Figure 2.3: Deformation of an arbitrary volume element with a representation of the displacement vectors between an initial state at time t_0 and a current state at time t_1

- $$\boxed{S_{ij} = \frac{1}{2}(u_{i,j} + u_{j,i})} \tag{2.9}$$

- $$\omega_{ij} = \frac{1}{2}(u_{j,i} - u_{i,j}) \tag{2.10}$$

S_{ij} is called the tensor of small deformations and ω_{ij} the tensor of small rotations.

Electrostatic: Electric field and displacement

To write the electrical part of the equations, the reference equations of electrostatic must be stated, including the hypotheses 3:

Hypothesis 3 *The following hypotheses are used to obtain the electric field equations:*

1. *The quasi-electrostatic approximation can be used because the phase velocity of acoustic waves is about five orders of magnitude smaller than the velocity of electromagnetic waves.*
2. *The magnetic effect is negligible compared to the electrical effect.*

Poisson's Law applied to the electric potential ϕ is defined as follows:

$$\nabla^2\phi = \frac{-\rho_e(M,t)}{\epsilon_0} \tag{2.11}$$

The Maxwell-Gauss law is defined such as :

$$\overrightarrow{\nabla} \cdot \overrightarrow{E} = \frac{\rho_e(M,t)}{\epsilon_0} \tag{2.12}$$

With $\phi(M,t)$ the electric potential at point M, $\rho_e(M,t)$ the charge density, $\overrightarrow{E}$ the electric field, $\epsilon_0 = \frac{10^{-7}}{4\pi c^2}$[F/m] the electric constant.

In Cartesian notation, the electric field is related to the electric potential as follows

$$\boxed{\overrightarrow{\nabla} \cdot \overrightarrow{E} = E_i = -\phi_{,i}} \tag{2.13}$$

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Electrical equation for an insulator

Panofsky [106] distinguishes the charge produced in an electric field between ρ_{true}, the true charge, and ρ_P, the polarization charge. This arbitrary distinction is therefore summarized as :

$$\rho_{true} + \rho_p = \rho \tag{2.14}$$

Poisson's law (2.11) becomes:

$$\nabla^2\phi = -\overrightarrow{\nabla} \cdot \overrightarrow{E} = \frac{-1}{\epsilon_0}(\rho_{true} + \rho_p) \tag{2.15}$$

The polarization volume charge corresponds to the variation of the dipole moment per unit volume:

$$\rho_p = -\overrightarrow{\nabla} \cdot \overrightarrow{P}. \tag{2.16}$$

Consequently, Poisson's law becomes, such as:

$$\overrightarrow{\nabla} \cdot \overrightarrow{E} = k_0^{-1}\rho_{total} = \epsilon_0^{-1}(\rho_{true} - \overrightarrow{\nabla} \cdot \overrightarrow{P}) \tag{2.17}$$

The electric displacement vector $\overrightarrow{D}$ in $[C.m^{-2}]$ is defined as:

$$\overrightarrow{D} = \epsilon_0 \overrightarrow{E} + \overrightarrow{P} \tag{2.18}$$

Thus, with the aforementioned equations:

$$\overrightarrow{\nabla} \cdot \overrightarrow{D} = \overrightarrow{\nabla} \cdot (\epsilon_0 \overrightarrow{E} + \overrightarrow{P}) = \epsilon_0 \overrightarrow{\nabla} \cdot \overrightarrow{E} + \overrightarrow{\nabla} \cdot \overrightarrow{P} = \epsilon_0 \epsilon_0^{-1}(\rho_{true} - \overrightarrow{\nabla} \cdot \overrightarrow{P}) + \overrightarrow{\nabla} \cdot \overrightarrow{P} = \rho_{true} \tag{2.19}$$

In the case of an insulator, there are no electric charges, with $\rho_{true} = 0$. So the electrical equation for an insulator becomes:

$$\boxed{\overrightarrow{\nabla} \cdot \overrightarrow{D} = D_{i,i} = 0} \tag{2.20}$$

Piezoelectric Constitutive Equations

The first law of themodynamics for piezoelectric media is written such as:

$$\dot{U} = \tau_{ij}\dot{S_{ij}} + E_i\dot{D}_i \tag{2.21}$$

Tiersen [139] defines the electrical enthalpy H by the relation:

$$H = U - E_i D_i \tag{2.22}$$

By time derivation, the relationship becomes:

$$\dot{H} = \dot{U} - \dot{E}_i D_i - E_i \dot{D}_i = \tau_{ij}\dot{S_{ij}} + E_i\dot{D}_i - \dot{E}_i D_i - E_i\dot{D}_i = \tau_{ij}\dot{S_{ij}} - \dot{E}_i D_i \tag{2.23}$$

The enthalpy depends on the deformation tensor and the electric field: $H = H(S, E)$; hence, by time differentiation :

$$\dot{H} = \frac{\partial H}{\partial S_{ij}}\dot{S_{ij}} + \frac{\partial H}{\partial E_i}\dot{E}_i \tag{2.24}$$

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Combining the two equations (2.23) and (2.24):

$$\frac{\partial H}{\partial S_{ij}}\dot{S_{ij}} + \frac{\partial H}{\partial E_i}\dot{E_i} = T_{ij}\dot{S_{ij}} - \dot{E_i}D_i \iff T_{ij} = \frac{1}{2}(\frac{\partial H}{\partial S_{ij}} + \frac{\partial H}{\partial S_{ji}}) = \frac{\partial H}{\partial S_{ij}} \tag{2.25}$$

Where, $\dot{S_{ij}} = \dot{S_{ji}}$ and H is constructed such that $\frac{\partial H}{\partial S_{ij}} = \frac{\partial H}{\partial S_{ji}}$ and $D_i = -\frac{\partial H}{\partial E_i}$.
In linear theory and considering a homogeneous quadratic solution, H is defined such as:

$$H = \frac{1}{2}c^E_{ijkl}S_{ij}S_{kl} - e_{kij}E_kS_{ij} - \frac{1}{2}\epsilon^S_{ik}E_iE_k \tag{2.26}$$

Where, $c_{ijkl} = c_{ijlk} = c_{jikl} = c_{klij}$ are the elastic constant corresponding to the Young's modulus of the material at $E = 0$ $[N.m^{-2}]$ or [Pa], $e_{ikl} = e_{kij}$ are the piezoelectric constant, which relates electrical displacement to strain and compressive stress to electric field when the transducer is blocked ($S = 0$) $[C.m^{-2}]$, and $\epsilon^S_{ij} = \epsilon^S_{ji}$ are the relative permittivity (or dielectric constant) at zero strain ($S = 0$) corresponding to the dielectric permeability of the material.

The constitutive equations of piezoelectricity are :

$$\boxed{T_{ij} = \frac{\partial H}{\partial S_{ij}} = c^E_{ijkl}S_{kl} - e_{kij}E_k} \tag{2.27}$$

$$\boxed{D_i = -\frac{\partial H}{\partial E_i} = e_{ikl}S_{kl} + \epsilon^S_{ik}E_k} \tag{2.28}$$

Where i, j, k, l take the values 1,2,3. These equations define the multiphysics of the material, linking the electrical and mechanical components of the piezoelectric element. In this thesis application case, piezoelectric transducers are used in the context of bending mode so $i = 3$, $j = 3$, $k = 1$ and $l = 1$ in the equations, with index 3 being the polarization axes by convention (Figure 2.4).

$$T_3 = c^E_{13}S_1 - e_{13}E_3 \tag{2.29}$$

$$D_3 = e_{13}S_1 + \epsilon^S_{33}E_3 \tag{2.30}$$

Duplicate indices are truncated like $11 = 1$ and $33 = 3$ and constants are defined as $c_{13} = c_{31}$ and $e_{13} = e_{31}$.

The equations (2.27) and (2.28) can also be written in a dual version, and the equations (2.29) and (2.30) can be written such as:

$$S_1 = s^E_{13}T_3 + d_{13}E_3 \tag{2.31}$$

$$D_3 = d_{13}T_3 + \epsilon^T_{33}E_3 \tag{2.32}$$

Where $s^E_{13} = \frac{1}{c^E_{13}}$ is the compliance matrix at $E = 0$ in $[m^2.N]$, $d_{13} = e_{13}s^E_{13}$ is the constant charge in $[m.V^{-1}]$ and $\epsilon^S_{33} = \epsilon^T_{33}(1 - k^2)$.

The energy conversion efficiency between mechanical and electrical energy can be quantified by the electromechanical coupling factor k^2. Mostly estimated from experiments [104], it is defined as the ratio between the energy converted and the energy supplied.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

The coupling factor is also defined as a dimensionless ratio of material coefficients, such as:

$$k^2 = \frac{d_{13}^2}{\epsilon_{33}^S s_{13}^E} \tag{2.33}$$

The identity between the two definitions is verified only in the case of planar stress conditions, where $E_1 = E_2 = 0$ [83] and the demonstration can be done using the stress-strain cycle. It can also be deduced from the equations (2.31) and (2.32), such as:

$$\begin{aligned} D_3 &= d_{13}(\frac{1}{s_{13}^E} S_1 - \frac{d_{13}}{s_{13}^E} E_3) + \epsilon_{33}^T E_3 \\ &= \frac{d_{13}}{s_{13}^E} S_1 + (\epsilon_{33}^T - \frac{d_{13}^2}{s_{13}^E}) E_3 \\ &= e_{13} S_1 + \epsilon_{33}^T (1 - k^2) E_3 \end{aligned} \tag{2.34}$$

The electric charge q can be obtained by integrating this equation over the electrode such that:

$$q = \int_0^{L_p} \int_0^{b_p} D_3 dx_1 dx_2 = \int_0^{L_p} \int_0^{b_p} [e_{13} S_1 + \epsilon_{33}^T (1 - k^2) E_3] dx_1 dx_2 \tag{2.35}$$

Where L_p is the length $[m]$, b_p is the width $[m]$ and h_p is the thickness $[m]$ of the piezoelectric transducer. The equation is simplified using the following statements:

- The electric field is uniform along the piezoelectric transducer: $E_3 = \frac{V}{h_p}$
- The thickness h_p is small compared to the mid plane z_m of the main substrate (Bernoulli assumption): $S_1 = -z_m u''$, where u is the substrate deflection.

$$q = (1 - k^2) C_p^S V - e_{31} z_m b_p [\dot{u}(x_2) - u' d(x_1)] \tag{2.36}$$

The constitutive equations allow the description of the piezoelectric properties as a sensor and actuator. Each transducer is defined with its internal piezoelectic capacitance C_p [26], under the hypothesis of a constant electric field along the thickness [20], such as:

$$C_p^S = \epsilon_{33}^T \frac{b_p L_p}{h_p} \tag{2.37}$$

Modal coordinates are used on n modes, such that:

$$u(x) = \sum_{i=1}^{n} z_i \phi_i(x) = -e_{31} b_b z_m V \sum_{i=1}^{n} \frac{\phi_i(x)(\phi_i'(x_2) - \phi_i'(x_1))}{\mu_i (s^2 + \omega_i^2)} \tag{2.38}$$

The piezoelectric transducer itself is used to excite the structure to respond and μ_i is the modal mass. The fraction of modal strain energy in the piezoelectric transducer can be defined, such that:

$$\nu_i = \frac{c_{13}^E b_p h_p z_m^2 (\phi_i'(x_2) - \phi_i'(x_1))}{L_p \mu_i \omega_i} \tag{2.39}$$

The effective electromechanical coupling factor for mode i is therefore:

$$K_i^2 = \frac{k^2 \nu_i}{1 - k^2} \tag{2.40}$$

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

The equation (2.36) becomes

$$\frac{q}{V} = C_p^S(1-k^2)(1+\sum_{i=1}^{n}\frac{K_i^2}{1+s^2/\omega_i^2}) \tag{2.41}$$

Considering q/V transfert function, the short circuit condition ($V=0$) defined the poles as the natural frequencies ω_i while the open circuit condition ($Q=0$) defines the zeros as a the natural frequencies Ω_i and s the Laplace variable. The modal expansion is dominated by the contribution of mode i, under the assumption that Ω_i is close to ω_i, such that:

$$1+\frac{K_i^2}{1-\Omega_i^2/\omega_i^2} \cong 0 \Leftrightarrow K_i^2 = \frac{\Omega_i^2-\omega_i^2}{\omega_i^2} \tag{2.42}$$

Equation 2.42 assumes that the modes at $q=0$ and $V=0$ are identical, so there is little coupling.

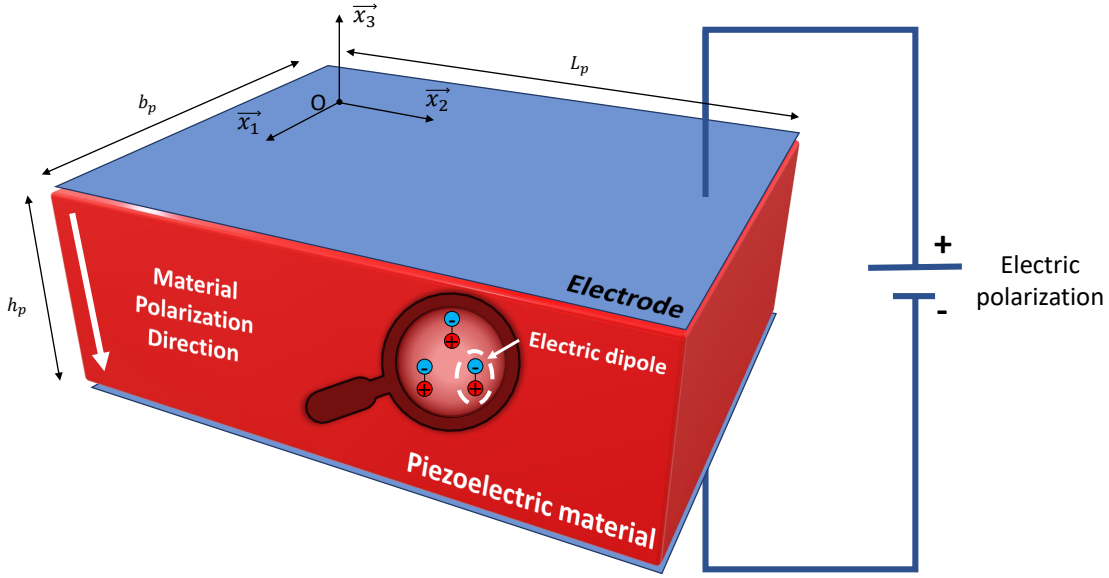


Figure 2.4: Architecture and orthotropic axes of a planar piezoelectric layer with electrodes being perpendicular to the polarization axis (the substrate is not represented).

2.1.1.3 From Continuous to Discrete Formulation: Variational Principle

Tiersen [139] describe the variational principle (Conservation of energy) for linear piezoelectric continuum, considering the internal energy balance with the Lagrangian L and the work of external forces W between two times $t0$ and t, such as:

$$\delta\int_{t0}^{t} L dt + \int_{t0}^{t}\delta W dt = 0 \tag{2.43}$$

$$\delta\int_{t0}^{t} dt \int_{Volume}[\frac{1}{2}\rho \dot{u_i}\dot{u_i} - H(S_{kl},E_k)]dVolume + \int_{t0}^{t} dt\int_{Surface}(t_k\delta u_k - \sigma\delta\phi)dSurface = 0 \tag{2.44}$$

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Where H the free enthalpy is the potential energy stored in mechanical and electrical form, f is the external mechanical loads (surface traction) and σ the electrical load (surface charge per unit area). A simplified formulation is obtained after applying initial and boundary conditions and using integration by part with time. The variational principle discetization lead to:

$$\int_{t0}^{t} dt[\delta u(t)\{(M_s+M_p)\ddot{u}(t)+(K_s+K_p)u(t)+K_{em}V(t)-bf(t)\}+\delta V(t)\{-K_{me}u(t)+C_pV(t)-q(t)\}] = 0 \tag{2.45}$$

Where M_s and M_p are the mass matrices, K_s and K_p are the stiffness matrices of the substrate and the piezoelectric elements, C_p is the dielectric constant, K_{em} and K_{me} are the piezoelectric coupling matrix (from electrical to mechanical and vice versa), bf is the mechanical load, q is the electrical load and V is the electrical voltage between the transducer electrodes. Since δu and δv are independent variables, the following system of equations is obtained:

$$\begin{array}{l} \delta u^T : \left\{ M\{\ddot{u}\} + J\{\dot{u}\} + K\{u\} + K_{\text{me}}\{V\} = b\{F\} \right. \\ \delta v^T : \left\{ K_{\text{em}}\{u\} - C_p\{V\} = -\{q\} \right. \end{array} \tag{2.46}$$

Where M and K are the total mass and stiffness matrices including the substrate and the piezoelectric element. A damping matrix J can be introduced to represent system dissipation.

2.1.2 Control of Piezoelectric Structures

To control vibrations, piezoelectric transducers are used as integrated sensors or actuators in structures. The problem, from a control point of view, is to anticipate or plan the effect of a signal on the physical system. It follows a methodology divided into the following items:

- Objective definition: optimal control, servo-control, output regulation
- System analysis: model definition with classification of input and output signals
- Controller synthesis: active and passive control are two main approaches presented in this literature review.

To provide a visual representation on an existing system, the proposed methods are applied to a 1-D piezoelectric beam system described in Chapter 3.

2.1.2.1 Active Control

In active control strategies, the motion of a structure is modified by a control system using external energy [132]. It can be used to reduce vibrations induced by disturbances in systems. Depending on the system output, the aim is to input the correct signal so that the sum of the waves results in vibration cancellation.

The main advantage of the method is its performance in terms of precision and effectiveness. However, the complexity of the implementation can vary between the chosen algorithms. Inspired by the work of Rahmani [114] and Housner [50], Figure 2.5 provides an overview of the most common active control algorithms. References with control applications on smart structures using piezoelectric transducers are also available.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

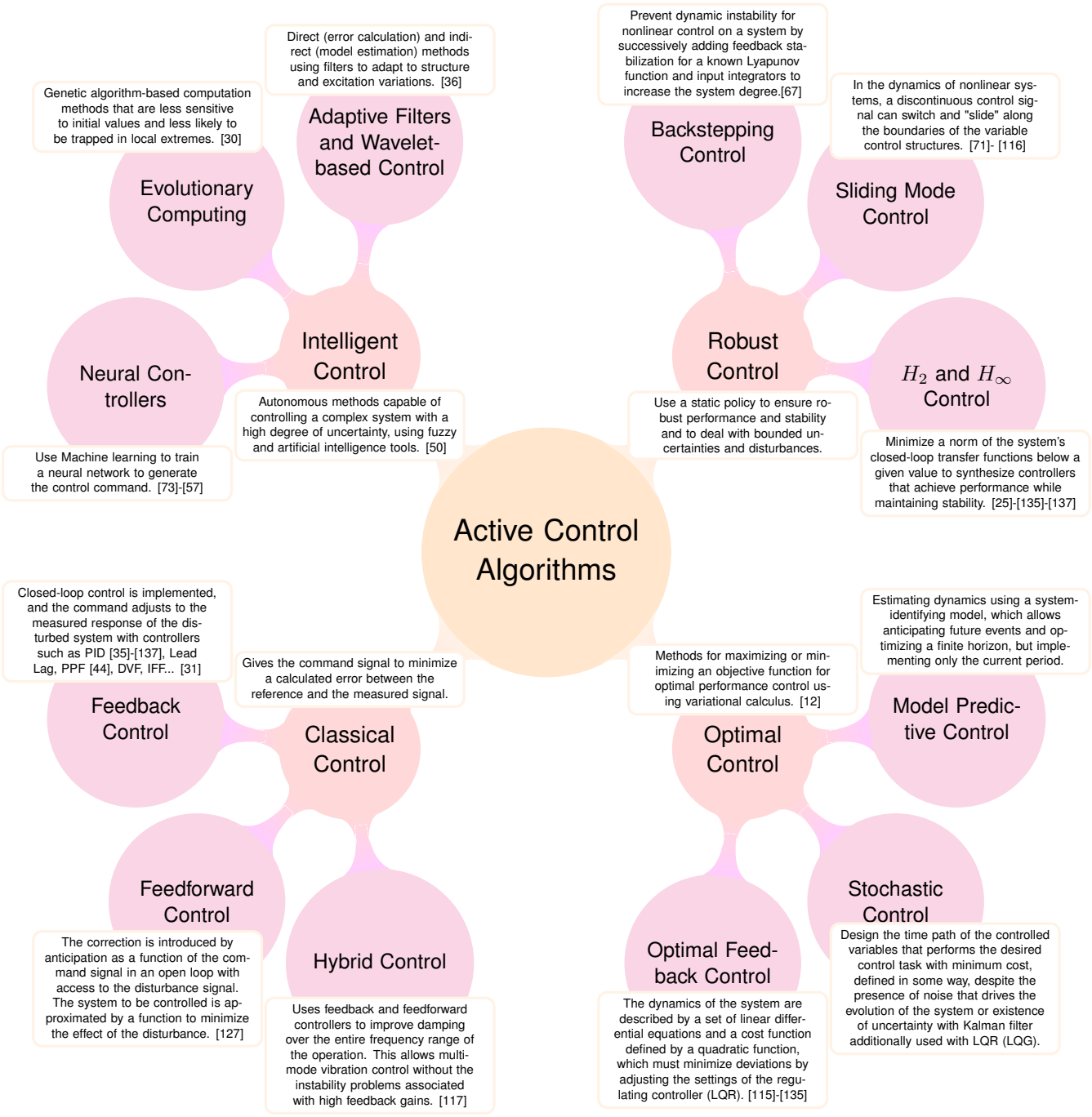


Figure 2.5: Active Control Methods Classification

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

A trade-off between complexity and efficiency must be defined when selecting the controller. System instability can also result from poorly designed control algorithms. Within the framework of existing algorithms, it is decided to focus on classical controllers, more specifically on feedback control methods, because of their simplicity of design and implementation. They are also efficient in the context of a linear structure.

Preumont [111] summarizes some feedback control methods and gives corresponding stability criteria with SISO (Single-Input-Single-Output) collocated transducers. To develop the following controller equations, the system $G(s)$ is defined as an open-loop transfer function, such as:

$$G(s) = \sum_{i=1}^{n} \frac{(b^T \phi_i)^2}{\mu_i (s^2 + 2\omega_i \xi_i s + \omega_i^2)} \tag{2.47}$$

Where s is the Laplace variable, n is the number of modes,$b^T \phi_i$ is the modal amplitude at the transducer location, μ_i the modal mass, ω_i are the natural frequencies of the system $[rad.sec^{-1}]$ and, ξ_i are the modal damping ratio. In the example case with the beam, the open-loop transfer function is taken between the two collocated piezoelectric transducers as introduced in Chapter 3.
Various feedback control laws $C(s)$ are applied to the system to control the vibrations as shown on Figure 2.6, where r is the target signal ($r = 0$ for the example), e is the error, $C(s)$ is the control law, v_d is the disturbance signal, v_c is the command signal, v is the input signal, y is the output signal. The close loop transfer function $G_{BF}(s)$ is written such as:

$$G_{BF}(s) = \frac{G(s)}{1 + G(s)C(s)} \tag{2.48}$$

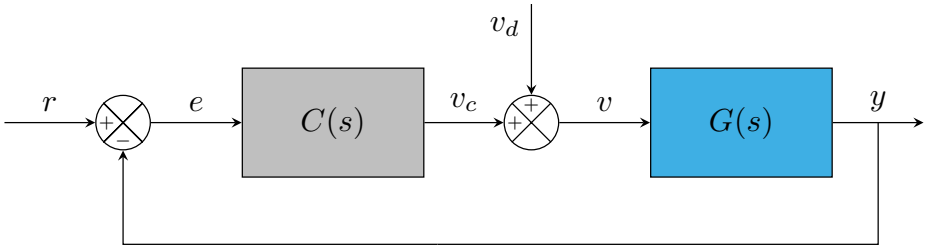


Figure 2.6: Feedback control law $C(s)$ applied on system $G(s)$

Stability and Robustness criteria

To apply control on a system, it is necessary to ensure stability. Various criteria can be used to quantify it. The first stability criterion is based on the poles of the system.

- Stability case: The controlled system is stable if all poles of the transfer function have a negative real part $Re(G_{BF}(s)) < 0$. Figure 2.7a provides a graphical analysis of the areas of stability and instability poles in the complex plane. The position of the pole is related to the damping through the imaginary axis angle ψ, such that $sin(\psi) \cong \xi$ (where $\xi << 1$). The more negative the real part of the pole is, the more damped the system is at the corresponding frequency.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

- Unstable case: With a positive real part of the pole, the damping would be negative ($\xi < 0$) and the system response will be amplified at each time step, leading to divergence, called system instability.

- Limit case: It occurs when one of the poles is null. Lyapunov stability can be discussed. The system is not damped at this frequency and gives an oscillatory response.

The second criterion was discovered by Nyquist in 1932 [102]:

- A closed-loop system is stable if and only if the Nyquist plot, described by increasing omega, surrounds the point (-1,0), the number of positive real poles of the function $1 + C(s)G(s)$. The point (-1,0) is called the critical point.

In most single input single output (SISO) cases, the transfer loop does not have poles with a positive real part, so the criterion is simplified and called the Bode criterion:

- A linear closed-loop system is stable if and only if the critical point lies to the left of the Nyquist plot (Figure 2.7b).

These criteria allow to introduce the concept of system robustness through the stability bounds.The phase margin Pm defines the amount of phase that can be added to the transfer loop before the system destabilizes, such that:

$$Pm = Arg[C(j\omega^*)G(j\omega^*)] + 180 \tag{2.49}$$

where ω^* is the pulsation such as: $|C(j\omega^*)G(j\omega^*)| = 1$.
The gain margin corresponds to the amount of gain that can be added to the transfert loop before the system destabilizes, such that:

$$Gm_{dB} = 20log_{10}(\frac{1}{|C(j\omega_0)G(j\omega_0)|}) \tag{2.50}$$

Where ω_0 is the pulsation in $[[rad.s^{-1}]]$ where the Nyquist plot intersects the negative real axis.

Both margins can be read from the Bode plot (Figure 2.7c):

- The phase margin Pm is the difference between the phase at the frequency where the transfer function module is zero and the horizontal asymptote at -180(deg).

- The gain margin Gm_{dB} is the difference between the module of the transfer function at the frequency where the phase is $-180(deg)$ and the horizontal asymptote at $0dB$.

Using the Nichols plot makes it easy to read the stability margins (Figure 2.7d). These criteria are limited to linear cases. In practice, reasonable margins are the gain margin $Gm_{dB} = 6dB$ and the phase margin $Pm = 45\,\text{deg}$, allowing a trade-off between control efficiency and system stability.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

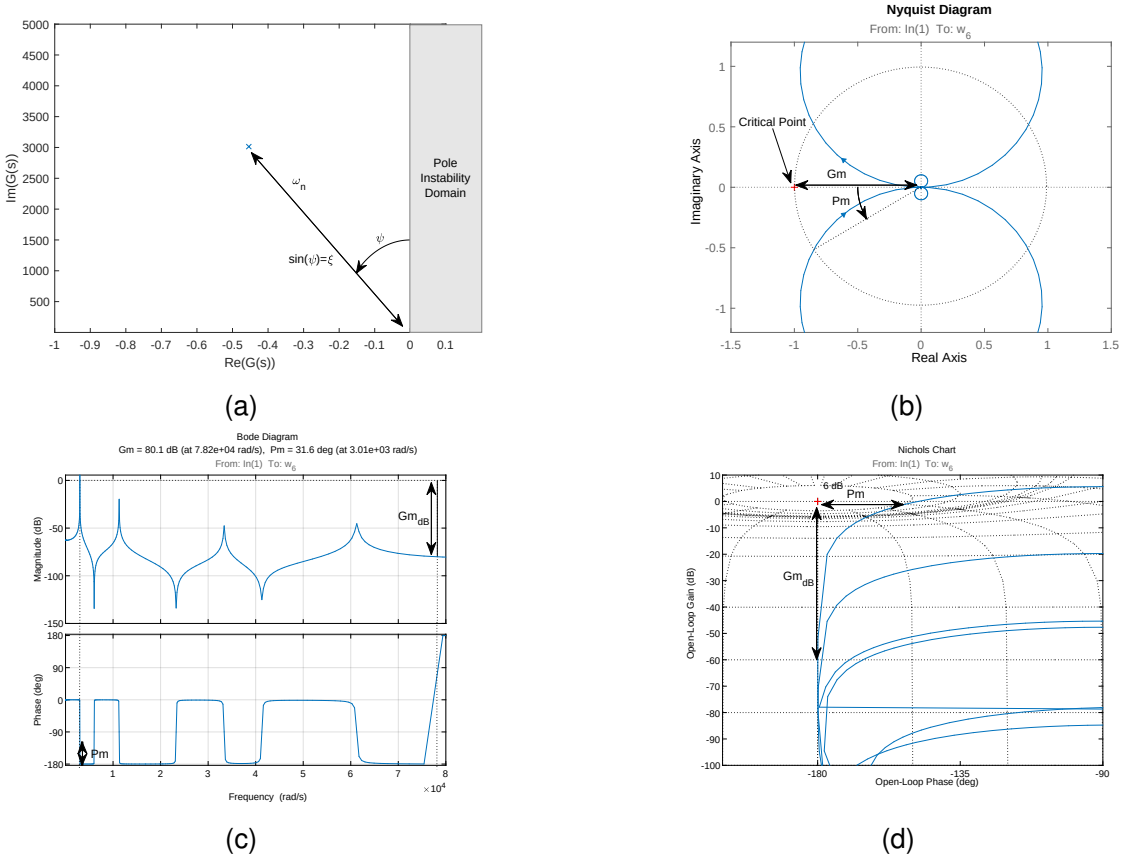


Figure 2.7: Visualization of the stability of the system $G_{FB}(s)$ where $C(s) = 40$ using: Complex Plane (a), Nyquist (b), Bode (c) and Nichols (d) plots.

Proportional Integrator Derivative (PID) Feedback:

As described in Borase works [11], the PID controller, widely used in industrial applications, is easy to implement on systems with a simple formulation, such that:

$$C(s) = K_p + K_i\frac{1}{s} + K_d s\frac{F_c}{s+F_c} = K_p(1 + \frac{1}{T_i s} + T_d s\frac{F_c}{s+F_c}) \tag{2.51}$$

Where K_p is the proportional gain, $K_i = \frac{K_p}{T_i}$ is the integrator gain, $K_d = K_p T_d$ is the derivative gain and F_c is the derivative filter frequency. Combinations of proportional, integrator and derivative controllers can be made depending on the control case. Their individual effects on the damping of the example system $G(s)$ are shown and discussed in the following lines. The most famous method for tuning a PID controller was introduced by Ziglers and Nichols [154] in 1942. The Chien, Hornes and Reswick (CHR) [17] in 1952 or Cohen-Coon in 1953 [19] tuning methods are also used in industrial applications.

A proportional controller is applied to the example system and its effect is shown in Figure 2.8:

- Figure 2.8a: The proportional term contributes to the control by adding a gain for error e through the closed loop and by modifying the static gain of the system
- Figure 2.8b: The closed loop static gain decreases, but the open loop static gain of the system increases as the controller gain increases, so the system stability decreases. The Gain (Gm) and Phase (Pm) margins are given in Table 2.3.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

- Figure 2.8c and 2.8d, orange lines: System oscillation can increase because the system time response is not reduced when the controller gain is above the so-called "ultimate sensitivity" K_{pu} where $Gm = 0$. It is close to the system instability that occurs when the open loop gain margin is negative $Gm < 0$.

- Figure 2.8c and 2.8d, purple lines: Ziegler and Nichols recommend finding this ultimate sensitivity gain and taking half the value to tune an efficient proportional gain $K_{pe} = 0.5K_u$. This chosen value is a compromise between stability and amplitude ratio.

In the example, note that the instabilities of the control system come from high frequencies, above the range of interest.

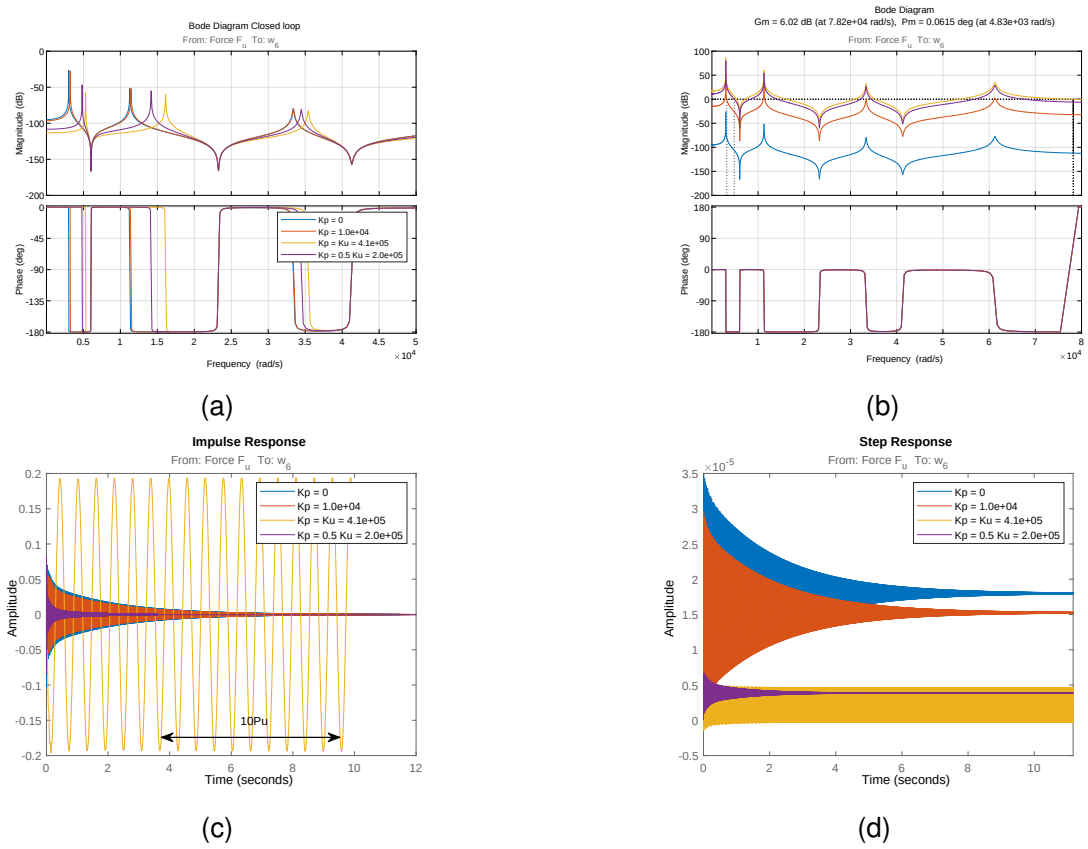


Figure 2.8: Closed-loop Bode (a), open-loop Bode (b), impulse (c), and step (d) responses of the $G(s)$ system with proportional feedback controllers for different gains.

Table 2.3: Gain and Phase margins of the open loop system $G(s)$ with Proportional controllers.

Kp	Gm (dB)	Frequency (rad/s)	Pm (deg)	Frequency (rad/s)
1.0e+4	3.2e+01	7.8e+04	0.1389	3.2e+03
Ku	1.5e-05	7.8e+04	0.0000	7.8e+04
Kpe	6.0e+00	7.8e+04	0.0615	4.8e+03

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

The effect of a Proportional Integrator (PI) controller on the system under study is shown in Figure 2.9 with different controller gains:

- Figure 2.9d: The integrator reduces the offset of the error e over time as an "automatic reset". It is defined by $T_i = \frac{1}{K_i}$, called the integral action time. To increase the accuracy, T_i must be increased so that the integration order is higher than the static error order.

- Figure 2.9b: The integrator must be combined with other controllers because an integral action alone can lead to system instability by adding a $\frac{-\pi}{2}$ phase. The margins are given in Table 2.4. As expected, the margins are smaller with the PI controller tuning than with the P controller alone.

The Ziegler-Nichols tuning method uses the period of the oscillation P_u generated at the ultimate sensitivity K_u, such as $K_i = 0.54K_u/P_u$.

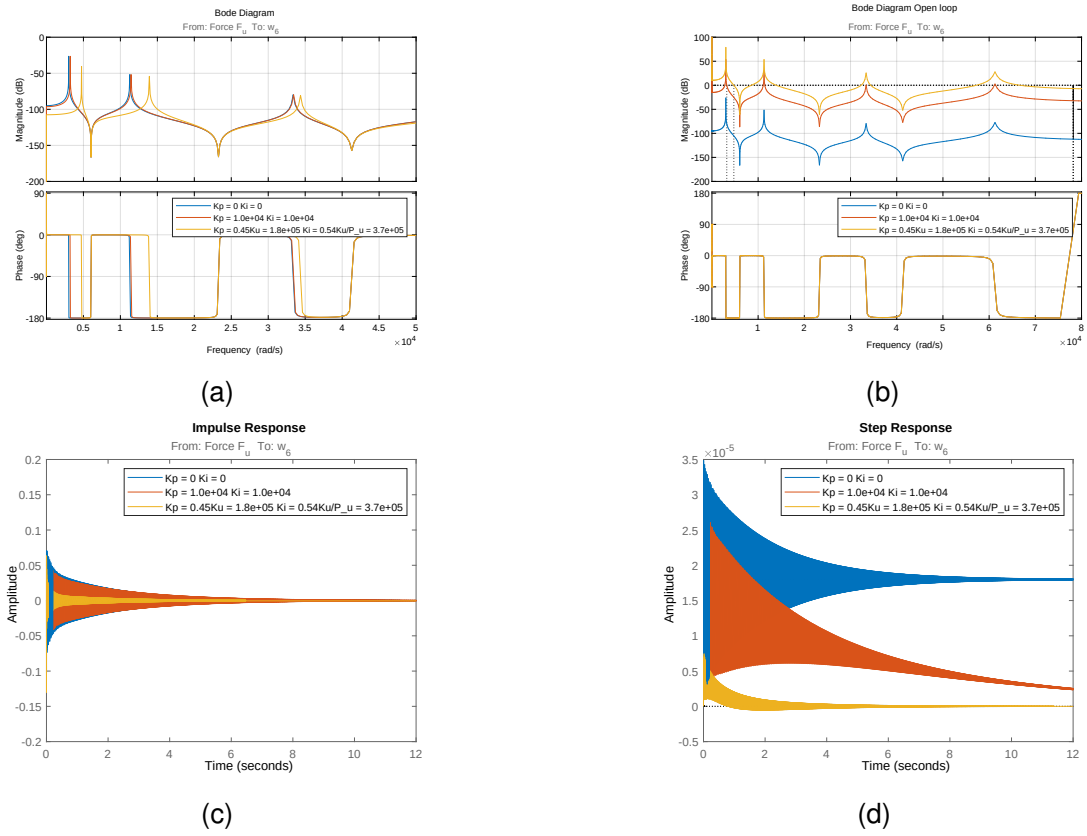


Figure 2.9: Closed-loop Bode (a), open-loop Bode (b), impulse (c), and step (d) responses of the $G(s)$ system with proportional integrator feedback controllers for different gains.

Table 2.4: Gain and Phase margins of the open loop system $G(s)$ with Proportional Integrator controllers.

Kp	Ki	Gm (dB)	Frequency (rad/s)	Pm (deg)	Frequency (rad/s)
1.0e+4	1.0e+4	3.2e+01	7.8e+04	0.12	3.2e+03
0.45Ku	$\frac{0.54Ku}{P_u}$	6.9e+00	7.8e+04	0.03	4.7e-03

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

The derivative gain monitors unusual variations of the error e by changing the output as a "pre-act" or an "anticipatory" controller. It is defined by $T_d = \frac{K_d}{K_p}$ the derivative action time.

A derivative controller is applied to the example system and its effect is shown in Figure 2.10. The use of a pure derivative controller is practically impossible due to causality. Instead, a pseudo derivative controller [101] with a time constant $T_f = \frac{1}{F_c}$ is used:

- Figure 2.10a and 2.10c: the derivative controller is tuned to damp the oscillations of the first mode.
- Figure 2.10b: As expected, using such a controller reduces the stability margin at frequencies above F_c (Table 2.5). The derivative action increases the phase of the controlled system and shifts the Nyquist plot away from the critical point, resulting in increased stability. The disadvantage is a decrease in static gain. It also induces high gain at high frequencies, which leads to noise amplification.
- Figure 2.10d: This form of controller does not correct the static error

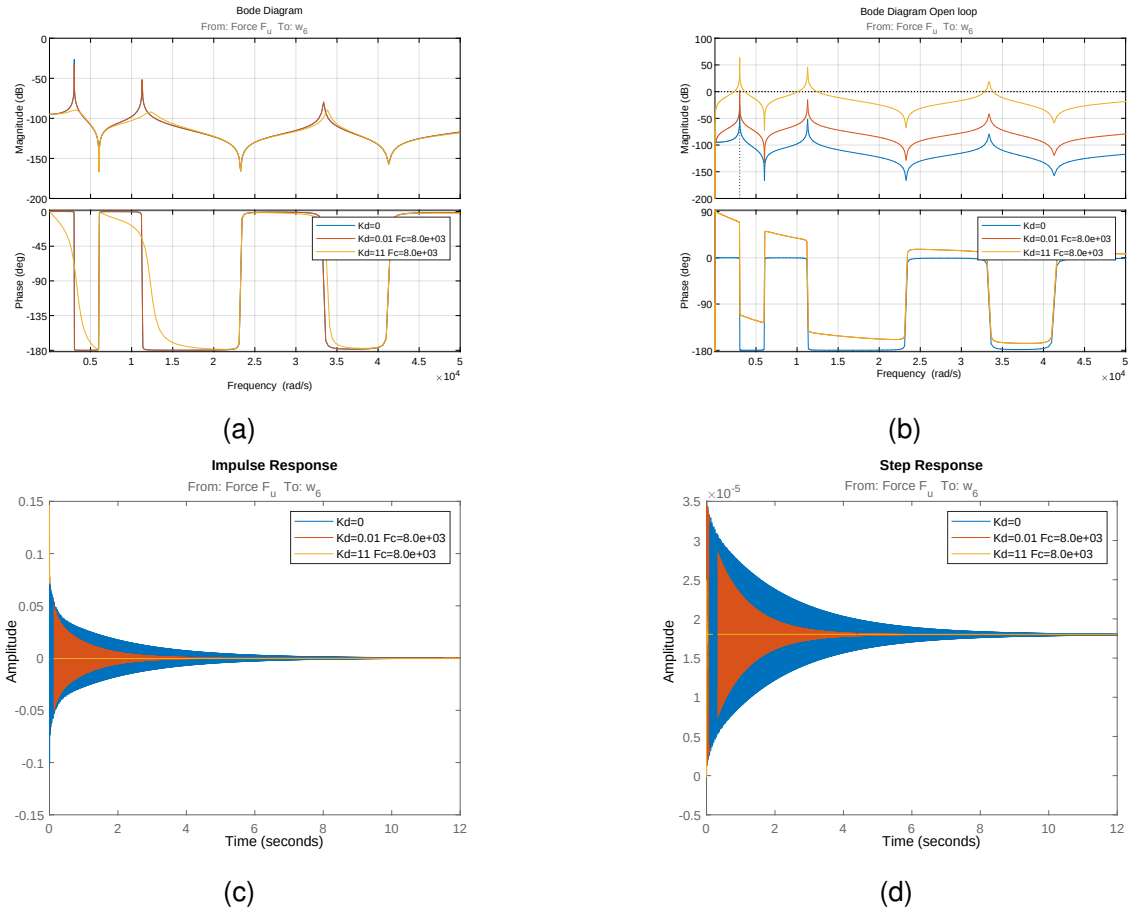


Figure 2.10: Closed-loop Bode (a), open-loop Bode (b), impulse (c), and step (d) responses of the $G(s)$ system with Derivative (D) feedback controllers for different gains.

The Ziegler-Nichols method of tuning uses the period of the oscillation P_u produced at the ultimate sensitivity K_u, such as $K_p = 0.6K_uP_u/8$. This method can not be used to tune K_d since a filter is added. Instead, a graphical method to tune the derivative

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

coefficient using root locus poles and zero variation with the gain K_d is shown in Figure 2.11. To damp the system, the real part of the first three poles must be as negative as possible without inducing a positive real part of the higher frequency poles.

Table 2.5: Gain and Phase margins of the open loop system $G(s)$ with Derivative (D) controllers.

	Kd	Gm (dB)	Frequency (rad/s)	Pm (deg)	Frequency (rad/s)
	1.0e-2	6.9e+00	9.9e+04	117	3.0e+03
Tuned Derivative	11	2.7e+00	9.9e+04	-2	1.0e+05

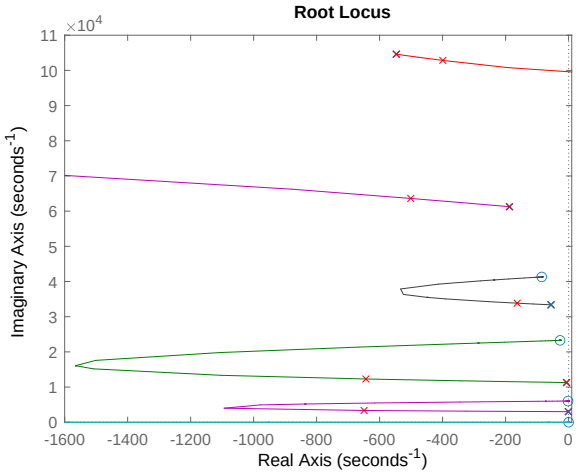


Figure 2.11: Root Locus of the system G with a Derivative controller and red (x) poles of the controlled system with $K_d = 21$.

In the literature, it is common to use PID controllers on piezoelectric cantilever beams. Khot et.al [64] implement it on a finite element model of a piezoelectric beam and uses iterative method to find optimal gain values by testing all possible ones. Jovanovic et.al [58] perform experimental verification showing the efficiency of Ziegler-Nichols tuning for different PID controllers on an intelligent cantilever beam damping with PZT actuators. Mathematical models are also developed to implement PID controllers on MIMO smart cantilever beam [70]. In this case of application, Ziegler-Nichols method to tune the PID is still efficient.

Lead-Lag Compensator:

Lead and lag compensators are designed to operate over a defined frequency range, so they can be specific and appropriate to the application. Analytically designed in 1976 by Wakeland [145] for the phase-lead compensation and refined by Mitchell [95] for the phase-lag compensation, the formulation is written as follows:

$$C(s) = K_l \frac{s + A}{\alpha s + A} \tag{2.52}$$

Where K_l is the gain, A is the frequency in $[rad.sec^{-1}]$, $\alpha < 1$ defines a lead compensator and $\alpha > 1$ defines a lag compensator.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

The effect of lead-lag compensators applied to the example system is shown in Figure 2.13. The tuned compensators and their margin values are listed in the table 2.6:

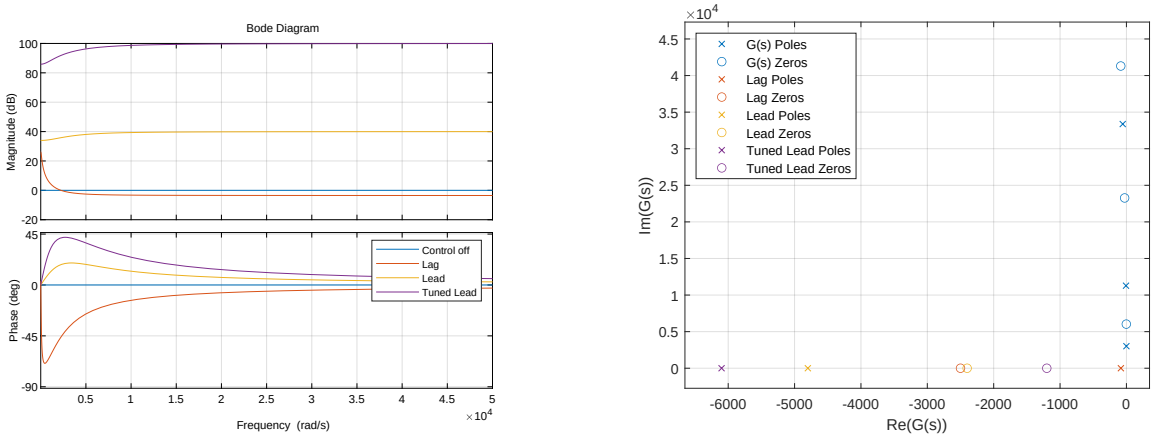
- Figure 2.13a: Lead and lag compensators affect damping at low frequencies.
- Figure 2.12a: The lead controller adds positive phase to the open loop transfer function (orange line) while the lag controller adds negative phase (red line).
- Figure 2.13b: The lead controller also improves the phase margin. The added phase is maximum and is equal to $\phi(\omega_m) = arcsin(\frac{\alpha-1}{1+\alpha})$ at $\omega_m = \frac{A}{\sqrt{\alpha}}$. (orange line). The lag controller reduced the phase margin, which can cause stability problems. The maximum loss phase is equal to $\phi(\omega_m) = arcsin(\frac{1-\alpha}{1+\alpha})$ at $\omega_m = \frac{A}{\sqrt{\alpha}}$ (red line).
- Figure 2.12b: The lead compensator pole and zero are designed so that the crossover frequency is between the two values (orange). The lag compensator is placed at low frequencies such as the poles and the zeros are below the crossover frequency of the uncontrolled system (red).
- Figure 2.13d: The aim of the lag compensetor is to exploit locally the integrator behavior of the controller and to increase the static accuracy.

Both controllers can be combined to operate in different frequency ranges.

In this thesis, a lead compensator is chosen for its stability improvement and is expressed in the following by :

$$C(s) = K_{ll}\frac{s+A}{s+B} \tag{2.53}$$

Where K_{ll} is the gain, A and B are the cutoff frequencies in $[rad.s^{-1}]$. To select a particular mode, the values of A and B must be selected on either side of the corresponding frequency. As shown in Figure 2.12a, the behavior of the tuned compensator (purple line) correspond to a lead (orange line) when $A < B$ and $K_{ll} = K_l/\alpha$.



(a) (b)

Figure 2.12: Bode (a) and Pole Zero (b) plots of Lead or Lag compensator for various gains.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

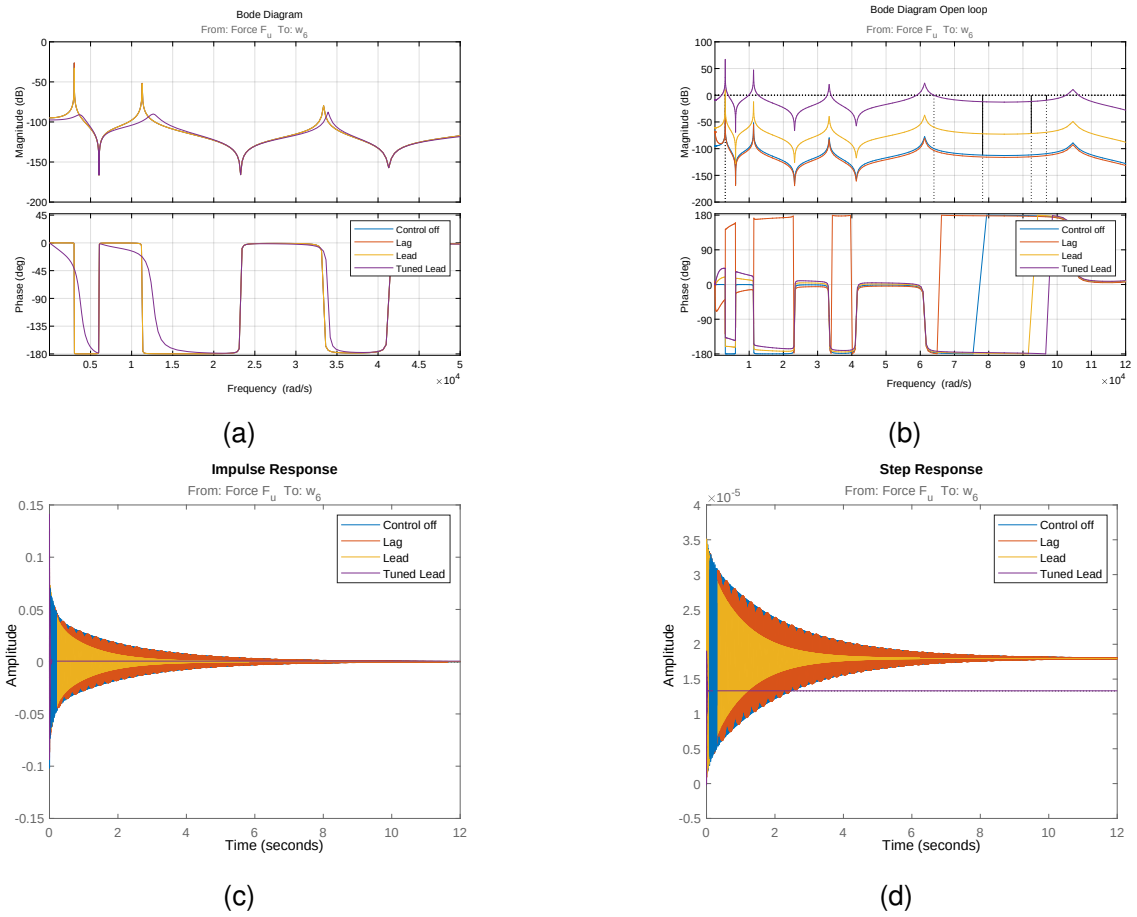


Figure 2.13: Closed-loop Bode (a), open-loop Bode (b), impulse (c), and step (d) responses of the $G(s)$ system with lead or lag compensators for different gains.

Table 2.6: Gain and Phase margins of the open loop system $G(s)$ with Lead or Lag compensators.

Type	K_{ll}	A	α	B	Gm (dB)	Frequency (rad/s)	Pm (deg)	Frequency (rad/s)
Lag	20	2.5e3	30	-	3.2e1	3.0e3	Inf	-
Lead	50	2.4e3	0.5	-	7.2e1	9.2e4	37	3.0e3
Tuned Lead	1e5	1.2e3	-	6.1e3	9.4	9.7e4	7.9	6.4e4

Finding the right lead-lag compensators usually comes after trial and error, even if a unique solution exists [148]. Kumar et al use Lead-lag compensators to account for phase lead and lag due to high-pass and low-pass filters coming from real-time controller boards to control vibrations of a flexible beam with PZT transducers [69]. The H_2 and H_∞ norm minimization are additionally selected with optimization algorithm using pattern search tools to tune the compensators and the added Positive Position Feedback (PPF) controller.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Positive Position Feedback (PPF):

Proposed by Goh and Caughey [38], the Positive Position Feedback (PPF) controller is a second-order filter such as:

$$C(s) = -K_{PPF}\frac{1}{s^2 + 2\xi_f\omega_f s + \omega_f^2} \tag{2.54}$$

Where K_{PPF} is the gain, ω_f is the filter frequency adapted to the target mode in $[rad.s^{-1}]$ and ξ_f is the damping ratio. This type of controller looks like an auxiliary system that acts as a tuned mass damper. It ensures control of low frequency modes while not destabilizing higher frequency modes. It is first imagined in case of large structure with high order and collocated transducers. The stability condition is defined as:

$$C(0)G(0) < 1 \tag{2.55}$$

To tune the PPF controller, first set ω_f to the mode frequency to be controlled, then set ξ_f usually between 0.5 and 0.7, and adjust the gain K_{ppf} with the root locus. As an example, PPF controllers are tuned to match the first and second frequency modes of the piezoelectric beam (Figure 2.14-2.16 and Table 2.7). Using this form of the controller adds a pair of poles to the system as it can be observed on the root locus (Figure 2.15a and 2.15b). The stability is increased by the phase added at low frequency (Figure 2.16b) with an increase of the static response (Figure 2.16a and 2.16d) due to the positive feedback gain.

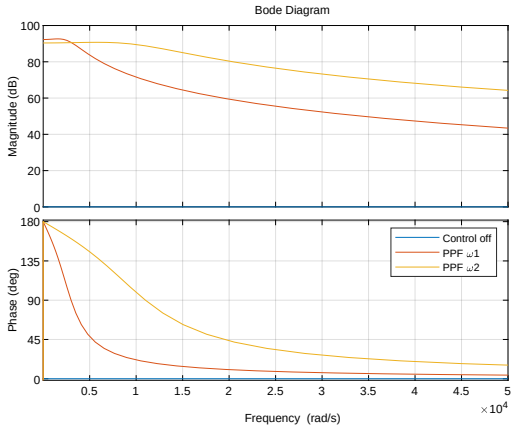
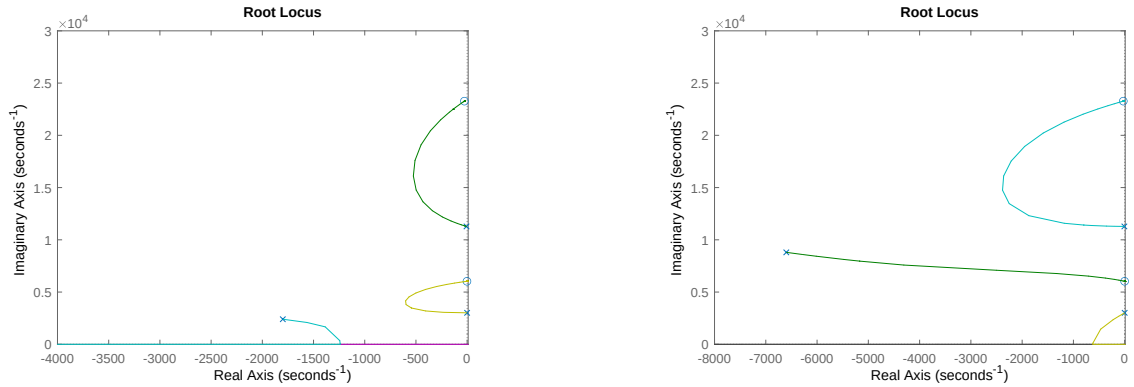


Figure 2.14: Bode plots of PPF controllers for different parameters

Table 2.7: Gain and Phase margins of the open-loop system $G(s)$ with Positive Position Feedback controllers.

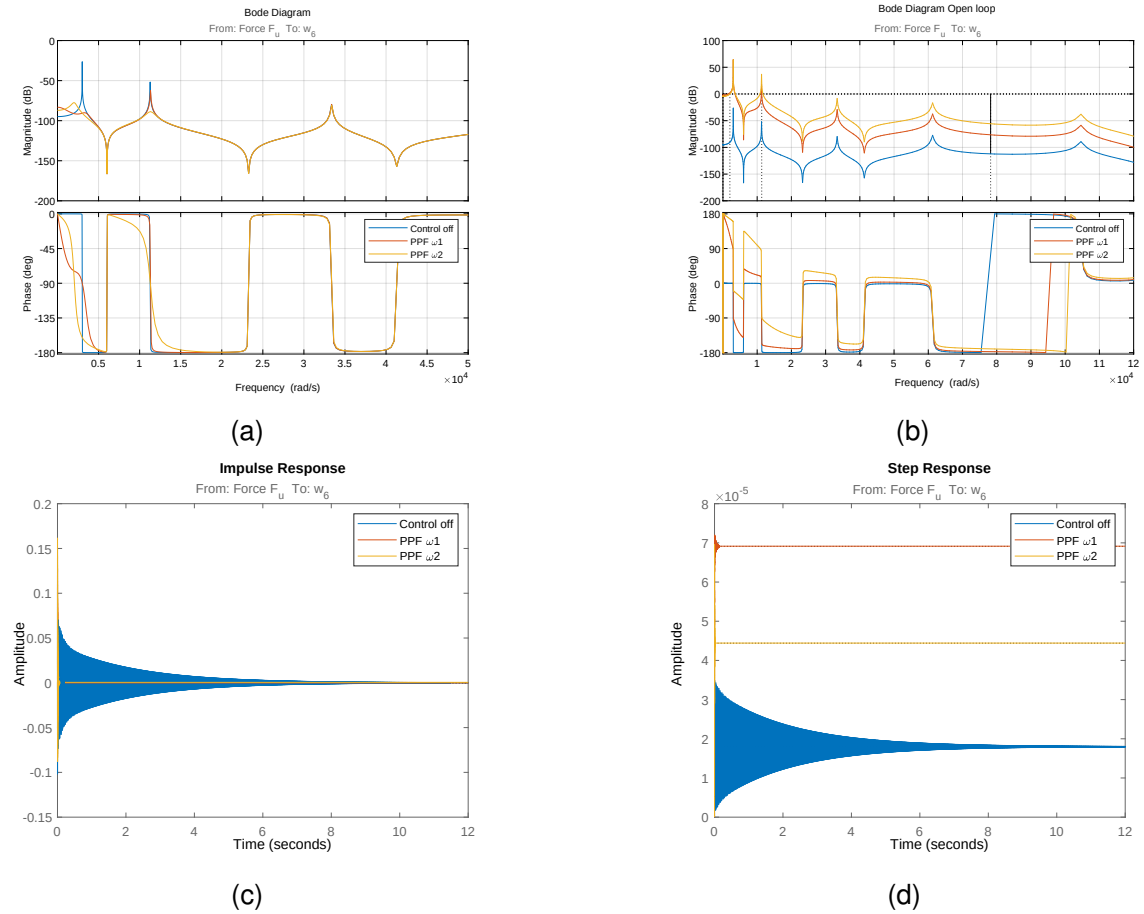
	K_{PPF}	ω_f	ξ_f	Gm (dB)	Frequency (rad/s)	Pm (deg)	Frequency (rad/s)
PPF ω_1	3.7e11	0.6	3e3	2.6	0	-26	1.1e4
PPF ω_2	4e12	0.6	11e3	4.5	0	-12	2.0e3

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés



(a) (b)

Figure 2.15: Root-locus plots of PPF controllers for different parameters: PPF tuned to target the first mode (a), and PPF tuned to target the second mode (b).



(a) (b)

(c) (d)

Figure 2.16: Closed-loop Bode (a), open-loop Bode (b), impulse (c), and step (d) responses of the $G(s)$ system with lead or lag feedback controllers for different gains.

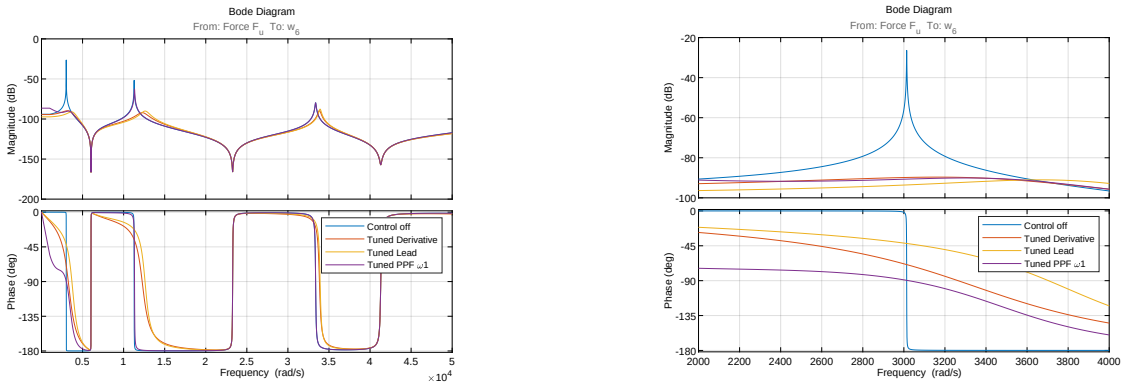
Li and al. [75] work on analytical and finite element approach for complex structure including piezoelectric beams. They use transfer function to represent system dynamics and implement PPF controller with efficient vibration damping. Lin and al. combine PPF controller with a fuzzy logic based observer for vibration damping but performances can still be improved [41]. To perform multimodal damping, PPF can also be used in

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

parallel [28]. Fanson et al. [27] conclude that in the case of high modal density at a given frequency range, a PPF filter should be used to control more than one mode. Lin [78] uses PPF with a time-switching method to perform multimodal damping when using a single piezoelectric actuator on a cantilever beam.

Impact of selected active control laws on the case study :

Figure 2.17 compares the effect of the controls previously shown and set to attenuate the first frequency of the structure. All controllers give similar performance, but the stability improvement at high frequencies increases with the use of Lead and PPF, so there are more parameters to tune.



(a) (b)

Figure 2.17: Closed-loop bode (a), with a zoom in frequency between 2000 and 4000 $rad.s^{-1}$(b) responses of the $G(s)$ system with the tuned studied feedback controllers.

In accordance with the literature, PID controllers seem reliable and easy to use in the context of vibration mitigation on piezoelectric structures. Depending on the expected performance, especially to improve high frequency stability, more complex controllers such as Lead-Lag or PPF may be required. Control laws need to be tuned considering performance, robustness and stability. Even though automatic tuning methods exist for tuning PID controllers, the most used method is still Ziegler-Nichols because of its simplicity. Also for lead-lag and PPF, tuning is mostly done manually after system modal analysis. Although active control methods are efficient, they require a power supply to inject energy into the system.

2.1.2.2 Shunt Control

Passive damping can be achieved by using piezoelectrics, so that no further energy is required to control the oscillations. Modes are damped by adjusting the transducer voltage with an electrical closed circuit between the transducer electrodes, or a "shunt". Various circuit models with analytical expressions of the optimized elements are available in the literature. The efficiency of this method depends on the amount of energy that can be transferred from the mechanical to the electrical domain thanks to the transducers. It is quantified by the modal coupling factor of the structure.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Modal coupling factor

As seen in Part 2.1.1.2, a piezoelectric structure is characterized by a modal effective electromechanical coupling factor from equation (2.42), such that:

$$k_i^2 = \frac{\Omega_{oc_i}^2 - \Omega_{sc_i}^2}{\Omega_{sc_i}^2} \tag{2.56}$$

Where Ω_{oc_i} are the natural frequencies of the structure when the piezoelectric elements have open-circuited electrodes and Ω_{sc_i} are the natural frequencies of the structure when the piezoelectric elements have short-circuited electrodes. Although these frequencies are usually very close to each other, structures can be designed to maximize the coupling coefficient. If the coupling coefficient is too low, the passive control will be negligible and will have no effect on the structure's behavior.

In order to perform a passive control on a structure, the piezoelectric equations are modified according to the boundary conditions of the transducer. Applied to the example system from part 3.1, the equation can be written for the short-circuit case where $V = 0$, such that:

$$\begin{gathered} M\{\ddot{w}\} + J\{\dot{w}\} + K\{w\} = F_{\mathrm{u2}}\{F\} \Rightarrow \\ \begin{cases} \dot{X} = \begin{bmatrix} 0 & \mathrm{Id} \\ -M^{-1}K & -M^{-1}J \end{bmatrix} X + \begin{bmatrix} 0 & 0 \\ 0 & M^{-1}F_{\mathrm{u2}} \end{bmatrix} U \\ Y = \begin{bmatrix} 0 & 0 \\ \mathrm{Id} & 0 \end{bmatrix} X + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} U \end{cases} \end{gathered} \tag{2.57}$$

The open circuit case is described by $q = 0$, such that:

$$\begin{gathered} M\{\ddot{w}\} + J\{\dot{w}\} + \left(K + K_{\mathrm{me}}K_{\mathrm{ee}}^{-1}K_{\mathrm{em}}\right)\{w\} = F_{\mathrm{u2}}\{F\} \Rightarrow \\ \begin{cases} \dot{X} = \begin{bmatrix} 0 & \mathrm{Id} \\ -M^{-1}K_{\mathrm{eq}} & -M^{-1}J \end{bmatrix} X + \begin{bmatrix} 0 & 0 \\ 0 & M^{-1}F_{\mathrm{u2}} \end{bmatrix} U \\ Y = \begin{bmatrix} 0 & 0 \\ \mathrm{Id} & 0 \end{bmatrix} X + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} U \end{cases} \end{gathered} \tag{2.58}$$

Where $K_{eq} = K + K_{\mathrm{me}}K_{\mathrm{ee}}^{-1}K_{\mathrm{em}}$ is the equivalent stiffness matrix. As a reminder, the short-circuit case corresponds to a zero resistance $R = 0$ between the two electrodes, while the open-circuit case corresponds to an infinite resistance $R = inf$.

Even if new methods exist to estimate the modal coupling coefficient [7], it is chosen in this thesis to use measurement approach with the system response as shown in Figure 2.18 with the example case.

The following lines show the effect of different electrical shunts on the example structure. Analytical tunings are taken from the literature and applied to damp the first mode of the structure.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

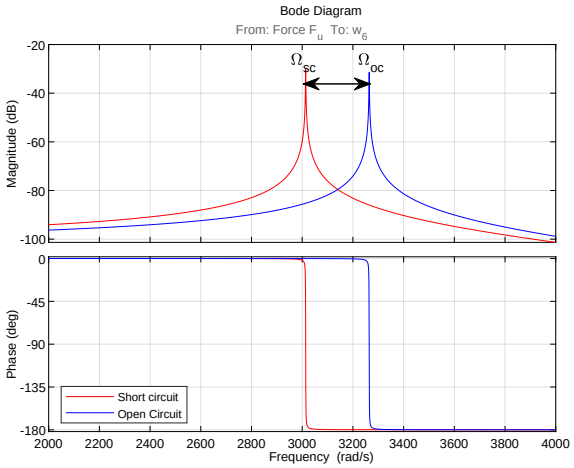


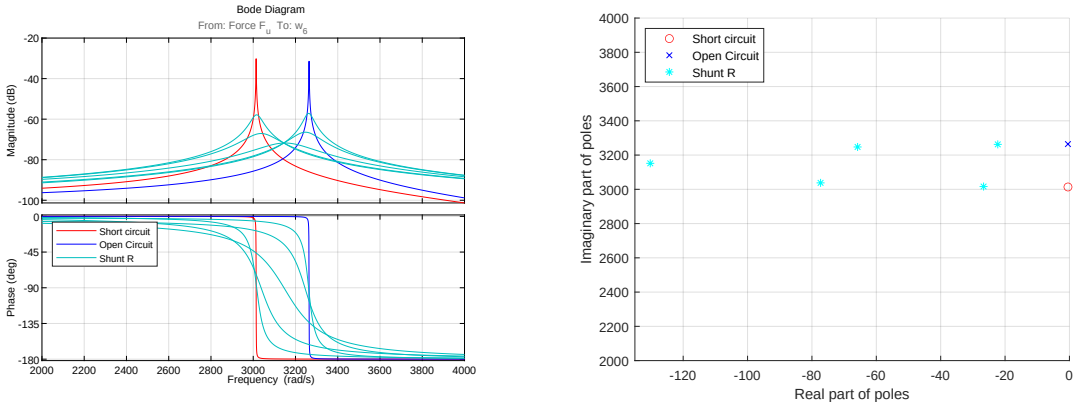
Figure 2.18: Bode response of the system $G(s)$ with (blue) short-circuit and (red) open-circuit piezoelectric electrodes zoom around the first mode.

Resistive shunt in serie

By adjusting the voltage across the electrodes of the piezoelectric transducers, the response of the structure is modified with frequency and damping variations. Figure 2.19 shows the effect of a resistive element inserted at the transducers electrodes on the example case, such as:

$$V = -Rsq \tag{2.59}$$

It defines what is called a resistive shunt, where R is the resistance of the shunt, s is the Laplace variable, and q is the electric charge.



(a) (b)

Figure 2.19: Bode of the system G(s) first mode (a) for a resistive shunt varying between 0 (short circuit) and infinity (open circuit) and (b) the corresponding root locus.

Preumont [112] defines the optimum resistive shunt as the system's maximum pole angle, which theoretically corresponds to the system's maximum damping (as seen in Part 2.1.2.1 Figure 2.7a):

$$\mathrm{R_{opt}} = \frac{1}{\gamma_{opt}\left[C_p\left(1-k_{31}^2\right)\right]} \tag{2.60}$$

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Where $\gamma_{\text{opt}} = \Omega_{oc}\sqrt{\frac{\Omega_{oc}}{\Omega_{sc}}}$ is the optimal feedback gain in the root locus graph. This formula is valid under the hypothesis of low coupling, but alternative methods can be used when high coupling occurs [98].

Resistive-inductive shunt in serie

In 1991, Hagood and von Foltow [43] defined two analytical formulations for the resistive-only and resistive-inductive shunt, defined such as:

$$V = -Rsq - Ls^2q \tag{2.61}$$

Where L is the inductance of the shunt. The optimal resistive-only shunt is defined, such that:

$$R_{opt} = \frac{\sqrt{1-k_{31}^2}}{\omega_n^E C_p} \tag{2.62}$$

Where ω_n^E is the natural frequency of the short-circuit system and $k_{13} = \frac{d_{31}^2}{s_{11}^E \epsilon_{33}^T}$ the electromechanical coupling factor.

For resistive-inductive shunt, the first method uses minimization of the transfer function maximum and defines optimal values such as:

$$\begin{cases} R_{optTF} = \frac{r_{opt_{TF}}}{C_p \omega_n^E} \\ L_{optTF} = \frac{1}{\omega_e^2 C_p} \end{cases} \tag{2.63}$$

Where $r_{opt_{TF}} = \sqrt{(2)}k_i/(1+k_i^2)$ is the dissipation parameter, $\omega_e = \delta_{opt_{TF}}\omega_n$ is the electrical resonance frequency, and $\delta_{opt_{TF}} = \sqrt{1+k_i^2}$ is the parameter for setting the piezoelectric shunt frequency.

The second method uses pole placement that maximizes the amplitude of the real part of the roots of the system and defines optimal values such that:

$$\begin{cases} R_{optPP} = \frac{r_{opt_{PP}}}{C_p \omega_n^E} \\ L_{optPP} = \frac{1}{\omega_e^2 C_p} \end{cases} \tag{2.64}$$

Where $r_{opt_{PP}} = 2\sqrt{k_i^2/(1+k_i^2)^3}$ is the dissipation parameter, $\omega_e = \delta_{opt_{PP}}\omega_n$ is the electrical resonance frequency and $\delta_{opt_{PP}} = 1 + k_i^2$ is the parameter for setting the piezoelectric shunt frequency. Figure 2.20 shows the significant effect of RL shunt on one degree of freedom (1-DOF) presented by Hagood and also on the first mode of the example system.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

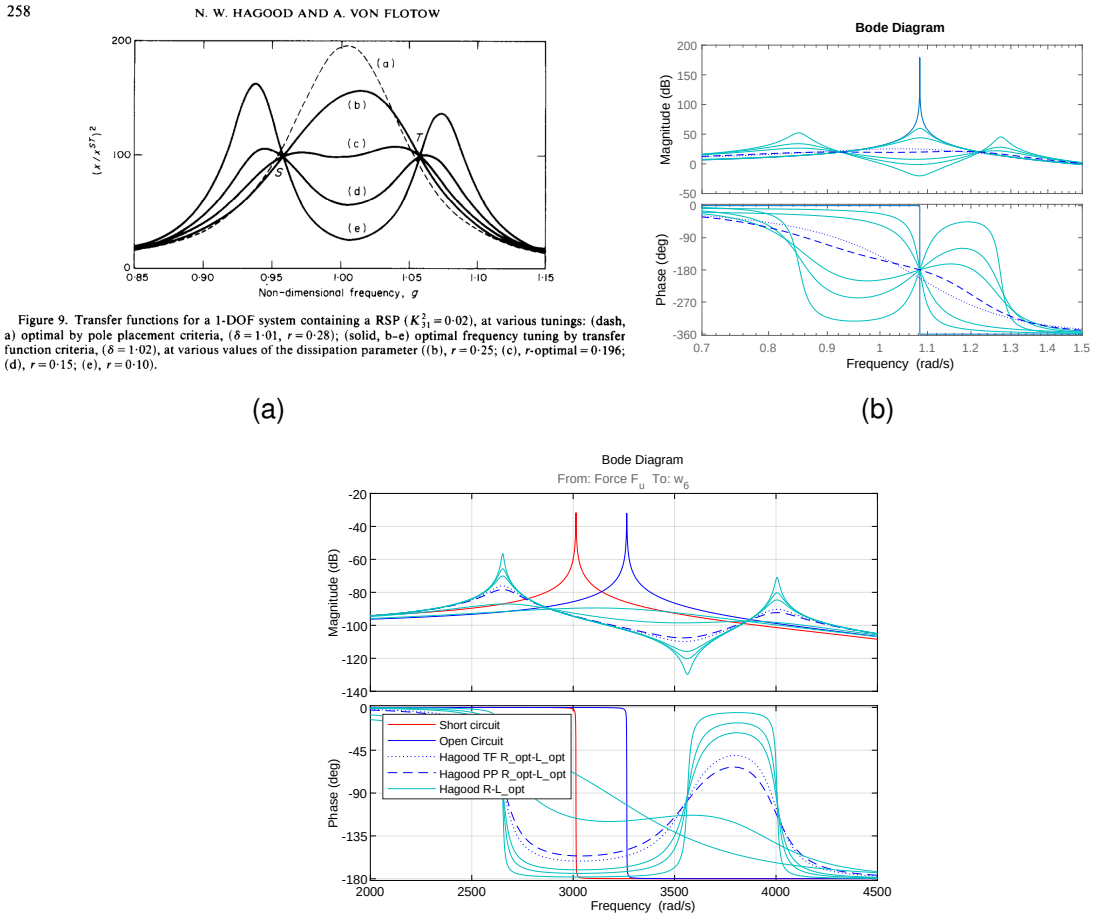


Figure 2.20: Hagood Transfer functions for a 1-DOF [43] compared to the system G(s) first mode for a resistive inductive shunt at various tuning (b) normalized with the short-circuit frequency and (c) not normalized.

Resistive and Negative Capacitance shunt

Collet et al [21] propose an improvement on Hagood's RL series shunt by using an RC shunt with negative capacitance, such that:

$$V = -Rsq - \frac{1}{C}q \tag{2.65}$$

Where C is the capacitance of the shunt tuned, such as:

$$C_{neg_{opt}} = -C_p \tag{2.66}$$

For a given system, there are three stability domains:

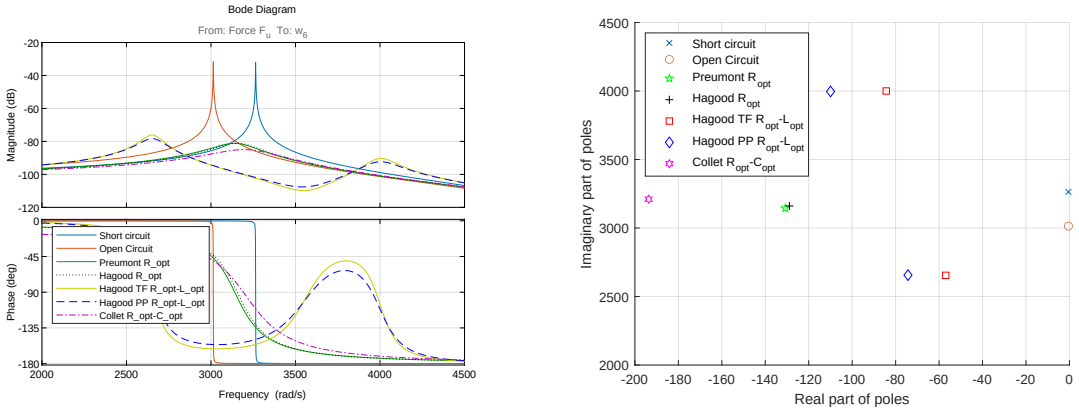
- $C_{value} <= -C_p^T$: Stable and softening domain
- $-C_p^T < C_{value} < -C_p^S$: Unstable domain
- $-C_p^S < C_{value}$: Stiffening stable domain

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Where C_p^T is the piezoelectric capacitance at zero stress and C_p^S is the piezoelectric capacitance at zero strain. This method can be considered as semi passive since external energy is necessary to create the negative capacitance circuit [10].

Impact of shunts on case study

Figure 2.21 highlights the passive attenuation of the piezoelectric patch induced by different optimal shunts and their effect on the first mode of the system example case.



(a) (b)

Figure 2.21: Bode (a) and Complex plane (b) of system G(s) response with various optimal shunt

For the resistive shunt, the theoretical optimum values obtained by Preumont and Hagood give similar, highly damped results for the beam's first bending mode. The RL series shunt is more attenuating than the Resistive shunt on the beam's first bending mode but it induces a peak splitting around the target frequency. The RC shunt is more effective in the frequency range, providing the best damping factor (Figure 2.21b). It is possible to use various combinations of resistive, inductive and capacitive shunts in series or in parallel as described by Berardengo et al. [8] on a cantilever beam coupled to a pair of piezoelectric patches. The use of negative capacitive shunts in series provides both stability and damping performance for vibration damping and, in some cases, resonance cancellation.

Literature highlights the possibility of using passive control on piezoelectric structure using various electrical shunt to do vibration damping. Optimal damping can be obtained with analytical formulation, but requires a relevant knowledge about the structure material and behavior. The following work is based on linear shunts only.

2.1.3 Optimization Algorithm

Optimization methods can be used to tune passive or active controller variables. Most optimization algorithms attempt to determine a set of input parameters for a function to be minimized or maximized. More than fifty exist but this thesis focuses on gradient descent, line-search and simplex algorithms.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

2.1.3.1 Gradient descent

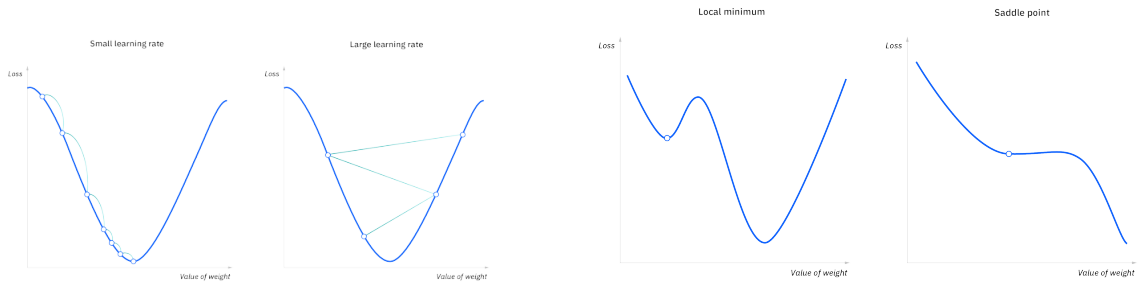
Proposed by Cauchy in 1847 [14], the gradient descent algorithm is an iterative method that uses the derivative of a cost function $f \in \mathbb{R}$, piecewise continuous, to find its minimum, called the convergence point. At the beginning, the derivative $\nabla f(x_k)$ is computed at a starting point $x_k = x_0 \in \mathbb{R}$, which gives the tangent direction of the function. The new point is chosen in the opposite direction of the slope with a step defined by a learning rate $\alpha \in \mathbb{R}$. The process continues until the tangent coefficient falls below the defined end values $\epsilon \in \mathbb{R}$, which are relatively close to zero. Algorithm 1 summarized the iterative process.

Algorithm 1 Gradient Descent Algorithm (1D example)

Require:

f ▷ Function to minimize
x_0 ▷ Initial Point
$\epsilon \geq 0$ ▷ End training conditions
α ▷ Learning Rate
$k \leftarrow 0$
while $||\nabla f(x_k)|| \leq \epsilon$ **do** ▷ Gradient ending condition check
    $x_{k+1} \leftarrow x_k - \alpha \nabla f(x_k)$ ▷ New point
    $k \leftarrow k + 1$
end while
x_k ▷ Point of convergence

For an efficient algorithm, the value of the learning rate is crucial. With large learning rate, divergence of the algorithm can occur because each time step will go far from the convergence point, whereas with small learning rate, convergence will take more time step to achieve (Figure 2.22a). This algorithm works for convex problems with global minimum. In the case of local minimum or saddle point, the success of the algorithm is highly dependent on the initial point chosen (Figure 2.22b). To avoid an infinite loop, breakpoint can be added to limit the number of loops.



(a) (b)

Figure 2.22: From IBM website: Illustration of gradient descent algorithm with various case: (a) small and large learning rate (b) local minimum and Saddle-point.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

2.1.3.2 Line-search

Line-search algorithm is similar to gradient descent but adjust its step size α at each iteration [100]. The step length is chosen to have a sufficient decrease condition with various methods like Newton or the coordinate descent methods. The coordinate descent does not require gradient calculation because variables are fixed in advance. At each iteration, the variable that reduces the most the objective function is selected. Algorithm 2 summarizes the iterative process using the coordinate descent method. A breakpoint can be added to limit the number of loops to avoid an infinite loop if the gradient end condition is not reached.

Algorithm 2 Line Search Algorithm (1D example)

Require:
f ▷ Function to minimize
x_0 ▷ Initial Point
$\epsilon \geq 0$ ▷ End training conditions
$\alpha \leftarrow [\alpha_1, ..., \alpha_N]$ ▷ N defined Learning Rate
$k \leftarrow 0$
while $||\nabla f(x_k)|| \leq \epsilon$ **do** ▷ Gradient ending condition check
  $p_k \leftarrow -\nabla f(x_k)$ ▷ Compute the descent direction
  $\alpha \leftarrow \alpha_1$
  $j \leftarrow 1$
  while $f(x_k) > f(x_k + \alpha_j p_k)$ **do** ▷ Find the biggest step size
    $x_{k+1} \leftarrow x_k + \alpha_j p_k$
    $j \leftarrow j + 1$
    if $j > N$ **then** ▷ Condition to Stop searching a learning rate too small
      Break
    end if
  end while
  $k \leftarrow k + 1$
end while
x_k ▷ Point of convergence

2.1.3.3 Downhill Simplex

The descent simplex method, also known as the Nelder-Mead method [99], is a nonlinear optimizer for multidimensional space functions $f(p)$, where $p \in \mathbb{R}^{\mathbb{N}}$ are the points and N is the dimension. The name simplex comes from the geometry field, where it denotes the simplest possible object for the given dimension: a line in 1-dimension, a triangle in 2-dimension and a tetrahedron in 3-dimension.

The algorithm starts with a simplex with $N + 1$ points defined in dimension N. An iterative process is used to find a point that minimizes f. At each iteration, f is computed at all simplex points to identify the minimum value p_1 and to compute the center of gravity p_o. Each point are order from the worst p_{N+1} to the best minimization p_1 of the function f. A reflection point p_r is defined as the symmetric point of p_1 from p_o. Then four cases can occur for the simplex: reflection, expansion, contraction or shrinkage. Those cases are detailed in Algorithm 3 with formulas.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Algorithm 3 Downhill Simplex Algorithm (N-dimension example)

Require:

f ▷ Function to minimize
$p \leftarrow [p_1, ..., p_{N+1}]$ ▷ Simplex Initialization with $N+1$ points
$\alpha \leftarrow 1$ ▷ Reflection coefficient (Standard value)
$\gamma \leftarrow 2$ ▷ Expansion coefficient (Standard value)
$\rho \leftarrow 1/2$ ▷ Contraction coefficient (Standard value)
$\sigma \leftarrow 1/2$ ▷ Shrink coefficient (Standard value)
$\epsilon \geq 0$ ▷ End training conditions
M ▷ Number of maximum iterations
$S_6 \leftarrow false$ ▷ Shrinkage Condition
while $std(p) \geq \epsilon$ or $k > M$ **do** ▷ Simplex ending condition check
  $f(p_1) \leq f(p_2) \leq ... \leq f(p_{N+1})$ ▷ Step 1: Order simplex points
  $p_o \leftarrow \frac{1}{N+1} \sum p$ ▷ Step 2: Calculate the simplex geometric center
  $p_r \leftarrow p_o + \alpha(p_o - p_{N+1})$ ▷ Calculate reflection point
  if $f(p_1) \leq f(p_r) \leq f(p_{N+1})$ **then** ▷ Step 3: Check Reflection
    $p_{N+1} \leftarrow p_r$
  else if $f(p_r) \leq f(p_1)$ **then** ▷ Step 4: Check Expansion
    $p_e \leftarrow p_o + \gamma(p_r - p_o)$ ▷ Calculate expansion point
    if $f(p_e) \leq f(p_r)$ **then**
      $p_{N+1} \leftarrow p_e$
    else
      $p_{N+1} \leftarrow p_r$
    end if
  else if $f(p_N) \leq f(p_r)$ **then** ▷ Step 5: Check Contraction
    if $f(p_r) \leq f(p_{N+1})$ **then**
      $p_c \leftarrow p_0 + \rho(p_r - p_o)$ ▷ The contraction point is outside the simplex
    else if $f(p_c) < f(p_r)$ **then**
      $p_{N+1} \leftarrow p_c$
    else
      $Step_6 \leftarrow true$
    end if
  else if $f(p_{N+1}) \leq f(pr)$ **then**
    $p_c \leftarrow p_0 + \rho(p_{N+1} - p_o)$ ▷ The contraction point is inside the simplex
    if $f(p_c) < f(p_{N+1})$ **then**
      $p_{N+1} \leftarrow p_c$
    else
      $S_6 \leftarrow true$
    end if
  end if
  if $S_6 == true$ **then** ▷ Step 6: Do Shrink
    $x \leftarrow x_1 + \sigma(x - x_1)$
    $S_6 \leftarrow false$
  end if
end while
p_o ▷ Point of convergence

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Reflection occurs when the reflected point gives a function value between the best $f(p_1)$ and the second worst $f(p_N)$. In this case, the original point p_{N+1} is replaced by the reflected point p_r. The expansion case occurs when the reflected point is the best of all simplex points. An expanded point p_e is calculated. If the function estimated at the expanded point is smaller than at the reflected point, the simplex is expanded. The contraction occurs inside or outside the simplex, depending on the function value estimated at the reflected point compared to the worst value. In both cases, a contracted point p_c is computed and compared to the second value, which can be p_r or p_{N+1} depending on the case. If the contracted point is better, simplex contraction occurs. If not, shrinkage occurs and all points but the best are replaced. Figure 2.23 shows the simplex geometric variation according to cases over a two-dimensional space.

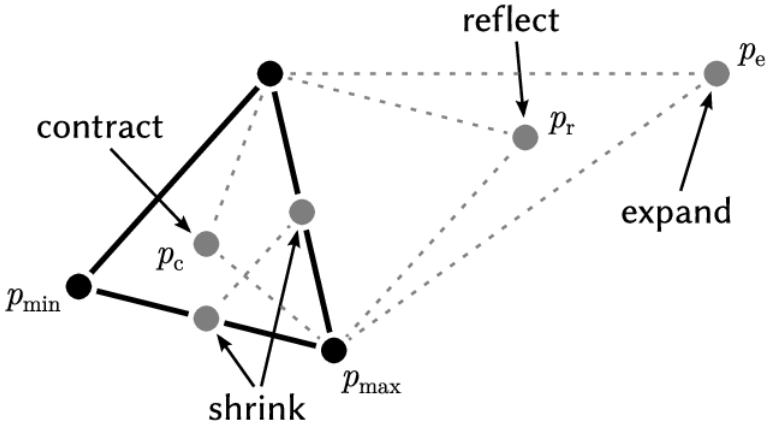


Figure 2.23: From Cheng [16]: Illustration of simplex geometric variation with cases

The halting criteria used in the original version of the algorithm compares the standard error of the simplex point with a preset value ϵ, such as:

$$std(p) = \sqrt{\frac{\sum(p_i - \bar{p})^2}{N+1}} \tag{2.67}$$

Where $\bar{p} = p_o$ is the average value of p.

As discussed by Lagarias et al [72], this method does not guarantee to find an optimal value in case of non-convex problems, but decrease significantly the objective value.

2.1.3.4 Genetic Algorithm

Based on the mechanism of natural selection, genetic algorithms were developed during the second half of the 20th century by John Holland's research team [49]. These methods can only find an optimum function in a finite space. The genetic algorithms are considered as blind, since they do not need any auxiliary information (derivative, ...) from the objective function. The algorithm starts with an initial population of points that evolve in parallel with probabilistic transition rules. The evolution process occurs over time with generations though elitism, mutation and crossover phonomenon in the population illustrated on Figure 2.24. These properties give a robust search for non-convex functions [39]. Algorithm 4 introduces the main steps of the methodology.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

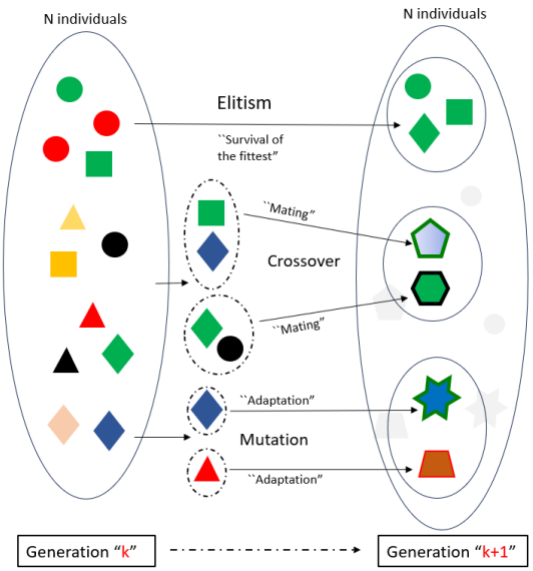


Figure 2.24: From Matanga [89]: Illustration of generation evolution.

Algorithm 4 From Katoch et al [60]: Classic Genetic Algorithm

Require:

f ▷ Function to minimize
N ▷ Population Size
N_e ▷ Elite size
N_c ▷ Crossover size
M ▷ Maximum Number of iterations
$P \leftarrow [p_{min}, p_{max}]$ ▷ Define the boundaries of the research space
$p^{(0)} \leftarrow p_1; ...; p_N$ ▷ Generate initial population randomly
$y^1 \leftarrow f(p_0)$ ▷ Compute the fitness values
$t \leftarrow 0$ ▷ Set iteration counter
while $t < M$ **do** ▷ GA ending condition check
  for $i = 1$ to N_e **do** ▷ Step 1: Select the N_e fittest (Elitism)
    $p^{(t+1)}{}_i \leftarrow min(f(p^{(t)}))$
  end for
  for $i = N_e + 1$ to $N_e + N_c/2$ **do** ▷ Step 2: Apply crossover operation on the $N_c/2$ selected pairs
    $[p_i^{(t+1)}, p_i^{(t+1)} + 1] \leftarrow crossover(p_{s1}^{(t)}, p_{s2}^{(t)})$ ▷ $p_{s1}^{(t)}$ and $p_{s2}^{(t)}$ are the selected parents
    $i \leftarrow i + 2$
  end for

  for $i = N_e + N_c/2 + 1$ to N **do** ▷ Step 3: Apply mutations to the elements that are left
    $p_i^{(t+1)} \leftarrow mutate(p_{s3}^{(t)})$ ▷ $p_{s3}^{(t)}$ is one of the left elements
  end for
$p^{(t+1)}$ ▷ Step 4: The newly created population replaces the old
  $t \leftarrow t + 1$ ▷ Increment the current iteration
end while
p_o ▷ Best solution point

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

These optimization algorithms are well known and have been used for a long time in engineering fields for their efficiency. Nowadays, gradient descent and line search algorithms are used for artificial intelligence approaches, especially for neural network training. Simplex and genetic algorithms are well known and used for optimization strategies. In the following chapters, these methods are used as indicators to compare with the efficiency of the proposed artificial intelligence methods.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

2.2 Theory of Machine Learning in AI

2.2.1 From Perceptron to Neural Network

2.2.1.1 History behind Neural Network

Connectionism vs Symbolism

There are two competing approaches to defining artificial intelligence through time: connectionism and symbolism (Figure 2.25).

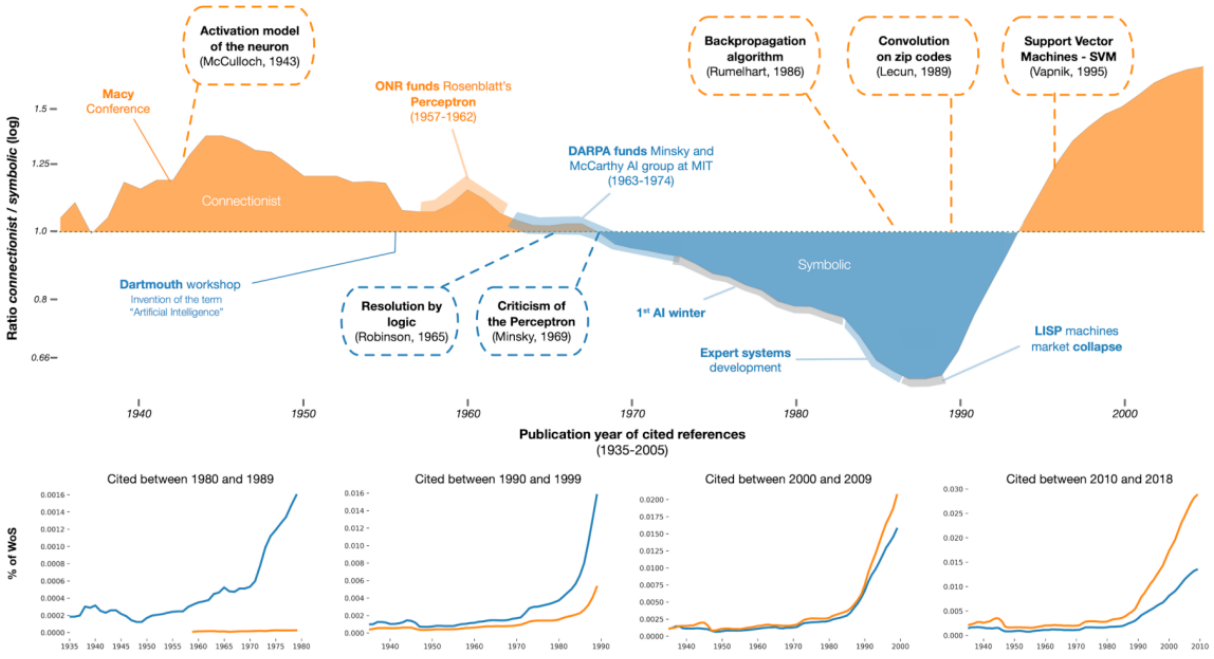


Figure 2.25: Connectionnist versus Symbolic Approaches Evolution between 1935 and 2018 by Cardon et al. [13]

Connectionism is an inductive approach. It attempts to model the brain by finding the laws that relate input and output data. In 1943, Culloch mathematically described the workings of brain activity using averages of logical propositions [91]. Based on the neuroscience of the time, Rosenblatt formulated the theory of the perceptron in 1958 [118]. This approach is criticized by Minsky and Papert [94] leading to the winter of connectionism by pointing out the slow updating process of networks. However in 1986, Rumelhart proposes a new method to train neural networks: Backpropagation [15]. This learning method becomes very popular because of its simplicity. An error function is defined according to a task to be accomplished, and a gradient descent is used to adjust the weight and bias of the neural network (NN), starting from the last layer to the first layer. Since then, the connectionist approach has become very popular, leading to the creation of many new areas of research.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Perceptron model

In 1958 Rosenblatt imagines the perception with the following hypothesis.

Hypothesis 4 *The probabilistic perceptron model is based on the following assumptions:*

1. *The construction of the neural network through physical connections is random.*
2. *The network's plasticity makes it adaptable to changing environments.*
3. *Exposure to broadly similar stimuli leads to a pathway to the same responding cells*
4. *Reinforcement leads to progress in connection training*
5. *Similarity in input signals induces similarity in output signals.*

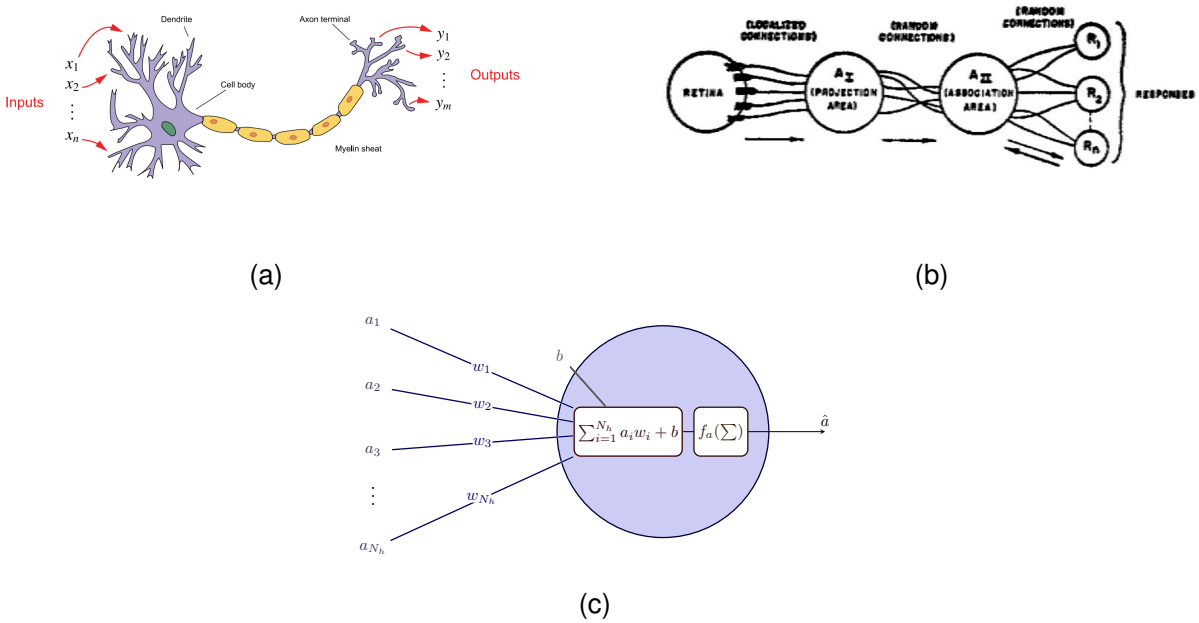


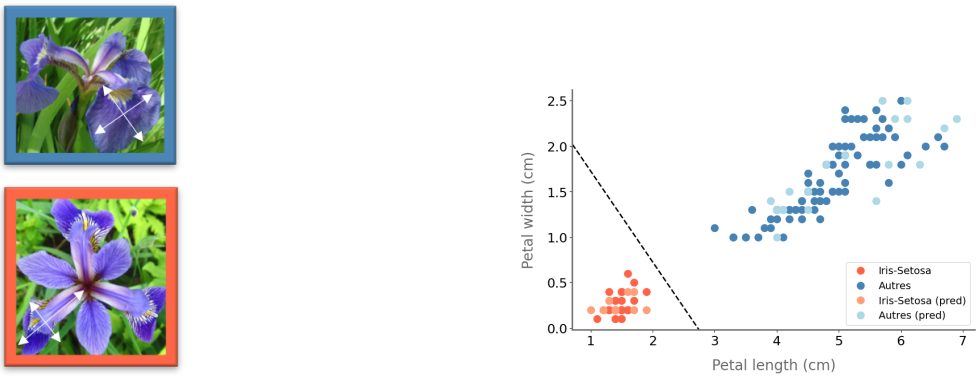
Figure 2.26: Organization of a perceptron: (a) Brain cell (b) Rosenblatt [118] and (c) modern representation

Figure 2.26 proposes a representation of a perceptron with a parallel between the original (Figure 2.26b) and the modern version (Figure 2.26c). Where $a_i \in \mathbb{R}$ are the input variables, $w_i \in \mathbb{R}$ are the weight, $b \in \mathbb{R}$ is the bias, $f \in \mathbb{R} \longrightarrow \mathbb{R}$ is the activation function and $\hat{a} \in \mathbb{R}$ is the output value. A perceptron can be used as a linear and binary classifier on data sets as for example on Figure 2.27. Here, the perceptron is trained to learn a pattern between the petal dimensions and the iris type. After training, it is able to identify the iris type by making a prediction according to the given input dimensions.

Neural networks Definition

Neural networks (NNs) are defined by assembling perceptrons, with layers differentiated into three parts: the input layer with $N_i \in \mathbb{N}$ input neurons, the hidden layers with

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés



(a) (b)

Figure 2.27: From FIDLE: Perceptron used as a classifier on (a) iris plants dataset from Fisher [29], with (dark color) train set and (light color) prediction set.

$N_{hu} \in \mathbb{N}$ hidden neurons per layer and $N_l \in \mathbb{N}$ layers, and the output layer with $N_o \in \mathbb{N}$ number of neurons. They are able to solve more complex real-world applications that are not only binary. Figure 2.28 is a schematic representation of a neural network (NN).

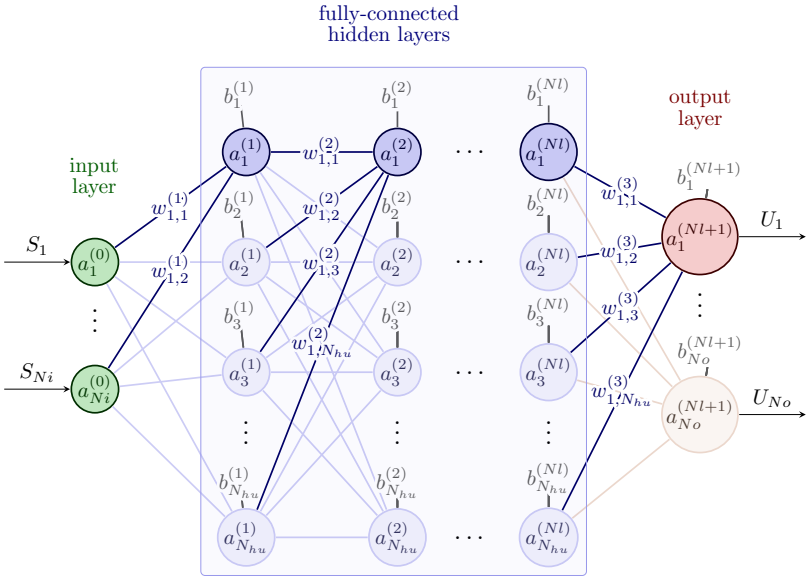


Figure 2.28: Neural Network schematic representation

Each layer $l \in \mathbb{N}$ is composed of independent neurons defined with a bias $b^{(l)} \in \mathbb{R}$ and an activation function $f_a^{(l)} \in \mathbb{R}^{n_{l-1}} \longrightarrow \mathbb{R}^{n_l}$. The connection between neurons in one layer and another is determined by a weighting parameter denoted as $w^{(l)} \in \mathbb{R}$. The output value $a_{n_l}^{(l)} \in \mathbb{R}$ of the neuron $n_l \in \mathbb{N}$ in the layer l is computed such that:

$$\begin{bmatrix} a_1^{(l)} \\ a_2^{(l)} \\ \vdots \\ a_{n_l}^{(l)} \end{bmatrix} = f_a^{(l)} \left(\begin{bmatrix} w_{1,1}^{(l)} & w_{1,2}^{(l)} & \dots & w_{1,n_{l-1}}^{(l)} \\ w_{2,1}^{(l)} & w_{2,2}^{(l)} & \dots & w_{2,n_{l-1}}^{(l)} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n_l,1}^{(l)} & w_{n_l,2}^{(l)} & \dots & w_{n_l,n_{l-1}}^{(l)} \end{bmatrix} \begin{bmatrix} a_1^{(l-1)} \\ a_2^{(l-1)} \\ \vdots \\ a_{n_{l-1}}^{(l-1)} \end{bmatrix} + \begin{bmatrix} b_1^{(l)} \\ b_2^{(l)} \\ \vdots \\ b_{n_l}^{(l)} \end{bmatrix} \right) \tag{2.68}$$

$$\mathbf{A}^{(l)} = f_a^{(l)}(\mathbf{W}^{(l)}\mathbf{A}^{(l-1)} + \mathbf{B}^{(l)}) \tag{2.69}$$

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

with $\mathbf{A}^{(l)} \in \mathbb{R}^{n_l}$ the output vector of the layer l, $n_l \in \mathbb{N}$ the number of neurons in the layer l, $\mathbf{W}^{(l)} \in \mathbb{R}^{n_l \times n_{l-1}}$ the weight matrix between the layers $l-1$ and l, $\mathbf{B}^{(l)} \in \mathbb{R}^{n_l}$ the bias vector of the layer l. The NN parameters to adjust in the training process are the bias and the weights of each layer.

Activation functions:

Activation function can vary according to the application case. For most case and for hidden layers, a ReLU [34] function f_a is chosen for neurons activation function, such as:

$$f_a(y) = \begin{cases} y & y \geq 0 \\ 0 & y < 0 \end{cases}, \quad y \in \mathbb{R} \tag{2.70}$$

For output layers other functions are used depending on outputs data types. The one used in this thesis are named Softplus and Hyperbolic tangent, and defined such as:

- Hyperbolic tangent: $f_a(y) = tanh(y)$
- Softplus: $f_a(y) = ln(1 + exp(y))$

The hyperbolic tangent activation function is centered at zero. Output values can be easily scaled or displayed as strongly negative, neutral or strongly positive.

2.2.1.2 The fields of AI

Today, AI is a term that encompasses many different fields. Figure 2.29. provides a non-exhaustive summary that represents existing categories. Two main branches are identified: Machine Learning and Expert Systems. The most famous tool used in artificial intelligence is deep learning, which is based on the use of neural networks. Many connections exist between the topic of artificial intelligence, but it can not be represented in a 2-dimensional plane. Expert systems mimic intelligence by following a set of rules or knowledge defined by a human expert. In another way, machine learning methods try to find these rules according to the input data. This data can be labeled or unlabeled, defining supervised or unsupervised learning. In both cases, the rules found can be stored by a trained neural network.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

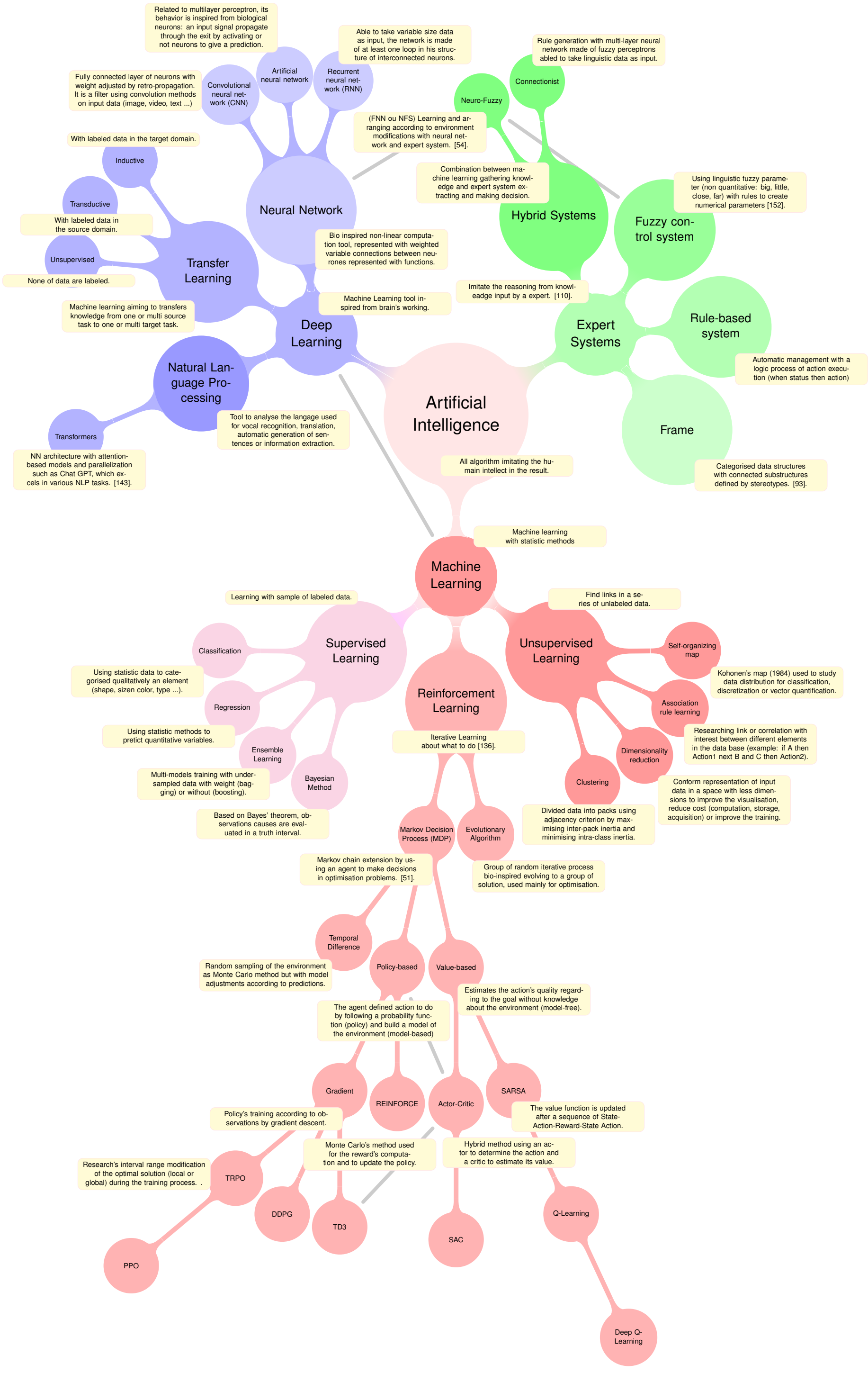


Figure 2.29: Artificial intelligence domains mapping

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

2.2.2 Reinforcement learning

2.2.2.1 Principe

Reinforcement learning is based on a numerical approach of iterative improvement without the need for a database. It differs from other machine learning approaches in that it defines a goal to be achieved. This objective is materialized by a mathematical function to be maximized, which is representative of the dynamics of the system under study. It is defined as a reward R as shown in the figure 2.30. Training is managed by an agent that takes input states S from the environment and chooses actions A_c to perform. As part of RL, Markov Decision Process with Policy-based Method defines the agent with an actor π and a critic V. The actor defines the policy (choice of action) to execute on the environment, while the critic estimates the expected reward from observations of the environment.
Richard S. Sutton and Andrew G. Barto [136] describe the mathematical and algorithmic foundations of this method in their book "Reinforcement Learning, An introduction". They define reinforcement learning as: **"learning what to do, how to associate situations with actions to maximize numerical reward"**, Chap1, 1.1 p1.

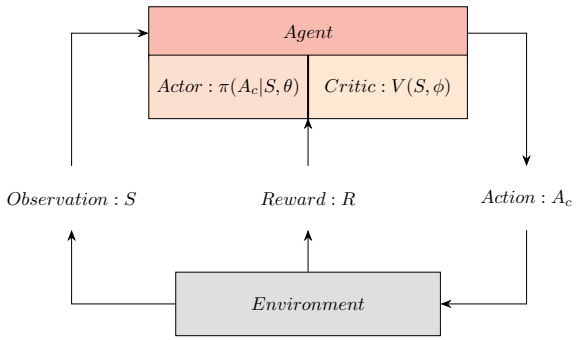


Figure 2.30: Reinforcement Learning algorithm: scheme

2.2.2.2 Deep Reinforcement Learning Algorithm

Deep Reinforcement Learning strategy uses Neural Networks (NNs) in the Agent as a non-linear function to optimize according to the reward definition. As seen in Part 2.2.1.2, many algorithms exist to train NNs. This thesis focuses on two of them: Trust Region Policy Optimization (TRPO) and Proximal Policy Optimization (PPO).

TRPO

Trust Region Policy Optimization (TRPO) [122] stores the policy (a course of action proposed) in a NN named the Actor π. It indicates which actions A_c to run on the environment according to the observations S, where the output layer is defined such as:

$$\pi(A|S,\theta) = \begin{cases} U_{i=\{1:\frac{N_o}{2}\}} = \mu = tanh(a_i^{(N_{l+1})}) \\ U_{i=\{\frac{N_o}{2}+1:N_o\}} = \sigma = tanh(ln(1+\exp(a_i^{(N_{l+1})}))) \end{cases} \tag{2.71}$$

Where θ is the Actor's NN weights and bias, $\mu \in \mathbb{R}^{\mathbb{N}\times}$ is the mean, $\sigma \in \mathbb{R}^{\mathbb{N}\times}$ is the standard deviation of the Gaussian probability distribution for each continuous action, $a_i^{(N_{l+1})}$ the

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

hidden layer output values and U_i are the $N_o \in \mathbb{N}$ NN outputs from equation (2.68). This model-based method builds a representation of the environment with a NN named the Critic V. It gives an estimation of an expected reward $R_V \in \mathbb{R}$ according to the observations S from the environment, such as:

$$V(S, \psi) = R_V \tag{2.72}$$

Where ψ is the critic's NN weights and bias. Main outlines of the TRPO algorithm updating process of ψ and θ are described in Algorithm 5:

Algorithm 5 TRPO Algorithm

Require: $N \geq 0$; R_{max} ▷ End training conditions
  while $t \leq N$ or $R_t \leq R_{max}$ **do**
    $A_c \leftarrow 0$ ▷ Initial Action
    $S_0 \leftarrow Env(A_c)$ ▷ Get Environment State
    $R_V \leftarrow V(S_0, \psi)$ ▷ Compute Critic output
    $A_c \leftarrow \pi(A_c|S_0, \theta)$ ▷ Compute Actor output
    $R \leftarrow Env(A_c)$ ▷ Get Environment Reward
    $\theta_{new} \leftarrow min(Loss_\pi(\theta))$ ▷ Minimize Critic Loss
    $\psi_{new} \leftarrow min(Loss_V(\psi))$ ▷ Minimize Actor Loss
    $\theta \leftarrow \theta_{new}$ ▷ Update Critic parameters
    $\psi \leftarrow \psi_{new}$ ▷ Update Actor parameters
    $t \leftarrow t + 1$
  end while
  $A_c \leftarrow Actor(S_0)$ ▷ Get end training Action

The agent's training is defined by two constraints: $N \in \mathbb{N}$, specifying the maximum number of episodes, and $R_{max} \in \mathbb{R}$ indicating the maximum reward target to be achieved. An Episode starts with the environment's initial state $S_0 = Env(A_c)$, defined here such as $A_c = 0$. The Critic estimates the expected reward R_V based on this initial state and the Actor chooses an Action A_c. The action is applied to the environment and Reward R can be computed to update θ and ψ parameters with Critic and Actor loss functions minimization using a line search algorithm (Part 2.1.3.2), such as:

$$Loss_V(\psi) = (R + \gamma V(S, \psi) - V(S, \psi))^2 \tag{2.73}$$

$$Loss_\pi(\theta_{new}) = -\frac{\pi(A|S, \theta_{new})}{\pi(A|S, \theta)}(R + \gamma V(S, \psi)) + w\frac{1}{2}\sum_{k=1}^{P} ln(2\pi.\exp(1).\sigma_k^2) \tag{2.74}$$

Where $P = N_O/2 \in \mathbb{N}$ is the number of output actions, $w \in \mathbb{R}$ is the Enthropy Loss Weigth, $\gamma \in \mathbb{R}$ is the Discount Factor, $\sigma_k \in \mathbb{R}$ is the standard deviation of the Gaussian probability distribution for each continuous action $k \in \mathbb{N}$ and exp() is the exponential function. Note that this formula is given in a simplified version where the batch size and the number of experiences per episode used are equal to 1.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

PPO

Proximal Policy Optimization (PPO) [121] is a simplified version of TRPO. The PPO algorithm has fewer hyperparameters and is therefore easier to tune and less computationally expensive than TRPO. The algorithm is mostly the same but the loss function to minimize with a gradient algorithm is such as:

$$Loss_{\pi}(\theta_{new}) = -min(\frac{\pi(A|S,\theta_{new})}{\pi(A|S,\theta)}(R+\gamma V(S,\psi)-V(S,\psi)), max(min(\frac{\pi(A|S,\theta_{new})}{\pi(A|S,\theta)}, 1+\epsilon), 1-\epsilon) + w\frac{1}{2}\sum_{k=1}^{P} ln(2\pi.\exp(1).\sigma_k^2)$$

Where ϵ is the clip factor. It should be noted that this formula is given in a simplified version in which the batch size and the number of experiences per episode used are 1.

Other policy gradient methods could be used, but TRPO and PPO are chosen because they are able to consider continuous observations and actions. Although TRPO requires more computation than PPO because it updates its policy within a trust region close to the current one, so it avoids performance degradation in most cases.

2.3 Piezoelectric structures and Artificial Intelligence

In the context of vibration control applications, several active methods have been developed that involve the integration of smart materials directly into structures. In general, these applications consider a limited number of transducers. However, the optimization of all control parameters becomes increasingly difficult as system dimensions and complexity grow, especially when physical uncertainties are taken into account. To solve this problem, researchers have recently explored intelligent control [76] in light of advances in computing hardware. Both fuzzy logic and artificial intelligence control are defined as intelligent control, but this part focuses only on AI.

2.3.1 Artificial Neural Network Controller

2.3.1.1 The beginning of ANN controllers:

The use of artificial intelligence in structure control appeared in the 1990s. Snyder and Tanaka [130] use a combination of a neural network and an adaptive LMS algorithm to remove disturbances from a structure in numerical simulation by feedforward control. Attention is also paid to the control of structures in response to environmental stresses (wind and earthquakes) [147].

Lee.G [73] introduced the use of NNs on a smart cantilever beam experiment for system identification, system state estimation or vibration control with the so-called Artificial Neural Network (ANN) controller.

Smyser et al [129] use a trained neural network to mimic a robust LQG/LTR controller on an analytical beam vibration model with integrated piezoelectric layers. The neural network is trained with a set of 5000 random input data, generated by the LQG/LTR controller with different initial conditions and forces applied to the beam. The vibration damping of the beam is faster with the neural controller than with the conventional

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

LQG/LTR controller, but the vibration wavelength is modified as the vibration amplitudes decrease.

Spencer et al [134] experimentally demonstrate the use of nonlinear adaptive neural networks on a beam structure, with sensor and piezoelectric actuators, for shape control and vibration suppression. Two collocated piezoelectric sensor-actuator pairs are used. On-line training of the neural network varies between 10 and 20 seconds for vibration minimization. As the sampling time is not specified, it is difficult to estimate the number of learning iterations for the simple neural network consisting of a layer of 10 neurons.

Valoor et al [142] use two neural networks in simulation to control the vibrations of an adaptive piezoelectric composite beam. A first identification network is pre-trained with 3000 data from an FE model. The second DRNN (Diagonal Recurrent Neural Network) is trained on-line for vibration reduction by controlling the terminal voltage of the piezoelectric elements.

Jha and Rower [57] conducted an experiment on a smart cantilever beam to test the usage of an ANN controller and its robustness against parametric uncertainties with multiple disturbance signals. They demonstrate the performance of a state feedback neuro-controller for vibration control. The controller's performance is quantified by a 13-20 dB reduction of the disturbances.

ANNs are also utilized as predictive models for feed-forward control to compensate for the hysteresis behavior of piezoelectric elements [77] or to attenuate disturbances on a cantilever beam [131] and other smart composite structures [97]. The works highlight a lack of consistency and predictability in performance due to the use of neural network controllers that are trained offline using Supervised Learning with a restrictive database, simulated by a model or measured on experimental setups.

In all those examples, a NN is used as a controller for vibration mitigation on numerical or experimental smart structure with a single feedback loop. The usage of a database is mandatory to train the NN and needs to be built according with the case. Building a database can be time consuming depending on the amount of data needed for the training and it can be biased by human selection.

2.3.1.2 Reinforcement Learning for ANN controllers:

According to recent advancements in Artificial Intelligence algorithms (Part 2.2.1.2) and hardware technology [74], new Machine Learning methodologies using NNs can be used to control systems like Reinforcement Learning (RL).

Considering Multiple Input Multiple Output (MIMO) systems, Unmanned Aerial Vehicles (UAV) have been recently controlled using NNs trained with Reinforcement Learning algorithm without the need for a database. RL controllers relevance compared to PID controller has been proven experimentally on trajectory tacking experiment on UAV [66].

Qiu [113] use a Reinforcement Learning (RL) controller to control a plate equipped with 5 piezoelectric transducers (4 for actuation and 1 for measurement) and minimize its response to sinusoidal excitation at resonant frequencies. Simulation results show that the RL controller is faster than a classical PD controller in damping the modal response. However, with the experimental setup, the RL controller has slower performance for transient vibrations but achieves similar performance after a few periods, depending on the frequency.

The performance of ANN controllers is dependent on the definition of the observation space, as demonstrated in the lift control experiment of a fixed-wing vehicle by Guerra-

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Langan [40]. In this experiment, various ANN, trained with Reinforcement Learning, are compared to a manually tuned PID controller. The ANNs are trained on a model which does not take into account noise and experimental uncertainties which explain why not all controllers are efficient experimentally. Also the training process requires an extensive amount of time and can not be performed on the experimental setup directly. It is mentioned that because ANNs controllers do not guarantee good performances, other non-linear approach could be used in this situation.

Indeed, using Reinforcement Learning on an experimental setup can be time-consuming depending on the complexity of the system, as shown in Howard's works [46]. Here, training one ANN requires 1000 episodes consisting of 200 timesteps of 0.05 seconds each, i.e. more than 2 hours.

All of these studies use an ANN trained with a reinforcement learning algorithm to provide the time domain control signal. This method is quite efficient in these application cases, but presents some time limitations in two ways. First, the training time required for NN is important. To perform it on an experimental setup requires robust and safe equipment that will not be disturbed or crashed during training. Second, the complexity of the neural network increases with the complexity of the system. The more actuators there are in the system, the more outputs there are in the NN. More computing time is required to compute the output, which introduces a delay in the feedback loop that can lead to a decrease in control performance or even experimental instability.

2.3.2 Neural Networks as Controller Optimizer on smart structures

More recent works have used reinforcement learning to set already existing and efficient controller parameters. For example, Khal [61] works with an active mass drive (AMD) system on structures stressed with seismic events. Reinforcement learning is used with online tuning to determine the gains of a fuzzy PD controller. Experimental tests allow to prove successful amplitude attenuation of the structural response amplitude in response to seismic perturbations.

In 2022, Pisarski et al [108] develop a switching control policy on a numerical Euler Bernoulli model of a multi-sensor beam monitor with semi-active control. A reinforcement learning method based on structure cost function minimization is implemented to adjust the controller parameters. By comparing with other methods, the results show the validation of the RL-based semi-active control method for transient vibration mitigation on a numerical model.

In 2024, Panda et al. [105] propose to use Reinforcement Learning algorithm with policy gradient based method to tune a P and a PI controller. The method's efficiency is illustrated on two numerical cases considering harmonic excitations: a quarter car model with active suspension system and an 8-story benchmark building.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

According with the recent literature, reinforcement learning methods are efficient to tune already existing forms of controllers in an autonomous way. It can also be extended to tune passive control laws as an electrical shunt on piezoelectric transducers. The use of this method for vibration mitigation on smart structures is still at the beginning especially in MIMO cases.

2.4 Conclusion: Position the subject within the state-of-the-art

New smart materials such as piezoelectric introduce transducers directly on the structure, creating light systems to control vibrations. The effect of the transducers on the structure must be maximum and taken into account during the construction process. It is then possible to implement efficient traditional control methods that exist with active or passive control. They are usually tuned by hand and must deal with stability limits. In the context of smart structure using piezoelectric transducers in a network, manual tuning becomes impossible by considering large structures. On the other hand, artificial intelligence tools are able to analyze a large amount of data using tools provided by neural networks. In addition, NNs are capable of memorizing patterns. These capabilities can be used to perform robust control for vibration mitigation on smart structures. Since this thesis works on various experimental test-bench, no database is available. It has been chosen to use reinforcement learning methods for its ability to train on a defined environment with a trial and error process. Neural networks are trained as optimizers to tune different controllers not only with active strategies but also with passive ones. The challenges are to adapt reinforcement learning tools to a vibration context, starting with a SISO system, then transferring to MIMO, and examining the robustness of the methods according to the system variations and perturbations.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Chapter 3

Introduction to the method: Deep Reinforcement Learning for passive control on a 1D element model.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

This chapter introduces the proposed method using the example case of a 1D finite element model of a cantilever beam coupled with two collocated transducers. It starts with the model properties with constitutive equations and boundary conditions. Then the deep reinforcement learning component and steps are described. The algorithm is run to tune the impedance of a shunt with the objective of damping the first mode vibration of the beam. The tuned shunt efficiency is compared with an analytical formulation from the literature. This reference system is also used in the previous chapter to illustrate the methods presented in the bibliography.
An open source code is available on a GIT repository for this chapter (Appendix A).

3.1 Smart cantilever beam model

3.1.1 Finite element model equations

Euler-Bernoulli theory is used to additionally model the beam with piezoelectric equations for the collocated transducers. A simple 1-dimensional finite element model of the beam can be defined, resulting in a coupled electromechanical system [80]. Figure 3.1 shows a representation of the beam and the system constants are defined in Table 3.1.

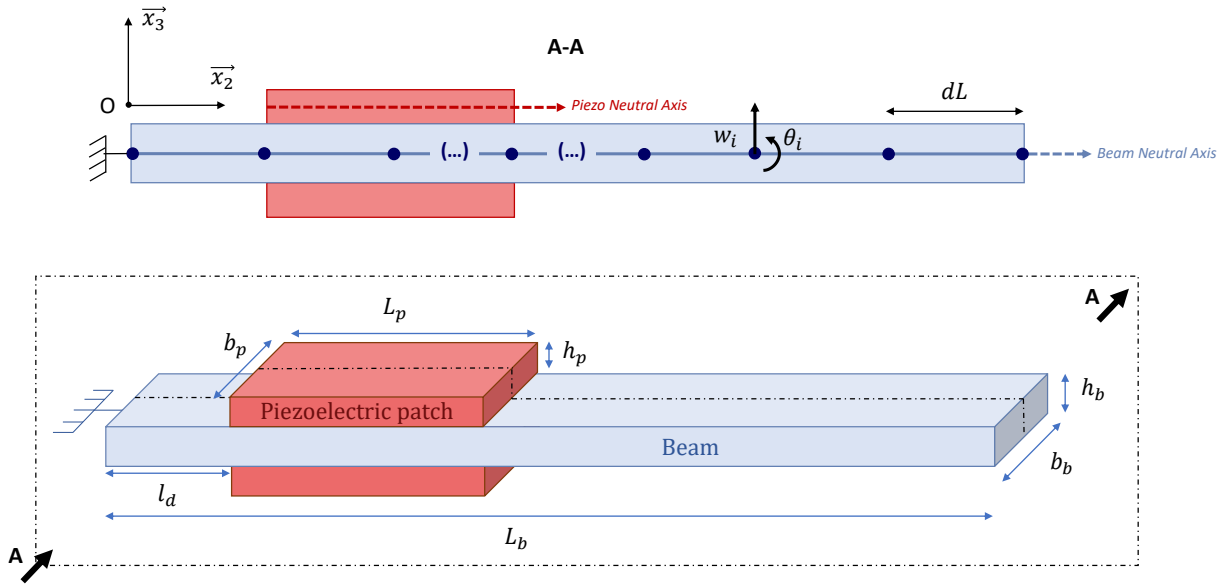


Figure 3.1: The reference system: Euler-Bernoulli cantilever beam with two collocated piezoelectric transducers.

Two types of elements must be considered: beam-only elements and piezo beam elements. The inertia along the neutral axis of the beam must be calculated, as well as the stiffness and density according to the case. The matrix elements are given in Table 3.2.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Table 3.1: Constant Definitions

	Variable	Beam	Piezoelectric Transducer	Unit
Dimension	Length	L_b	L_p	$[m]$
	Width	b_b	b_p	$[m]$
	Thickness	h_b	h_p	$[m]$
	Location	-	L_d	$[m]$
Material Properties	Young Modulus	E_b	$E_p = c_{13}^S$	$[Pa]$
	Density	ρ_b	ρ_p	$[kg.m^{-3}]$
	Coupling Constant	-	e_{13}	$[C.m^{-2}]$
	Relative Permittivity	-	ϵ_{33}^S	$[-]$

Table 3.2: Elementary matrix Definition

Type	Name	Formulation
Stiffness	K_{elem}	$\frac{\text{EI}}{\text{dL}^3}\begin{bmatrix} 12 & 6\text{dL} & -12 & 6\text{dL} \\ 6\text{dL} & 4\text{dL}^2 & -6\text{dL} & 2\text{dL}^2 \\ -12 & -6\text{dL} & 12 & -6\text{dL} \\ 6\text{dL} & 2\text{dL}^2 & -6\text{dL} & 4\text{dL}^2 \end{bmatrix}$
Mass	M_{elem}	$\frac{\rho\ h\ b\ \text{dL}}{420}\begin{bmatrix} 156 & 22\text{dL} & 54 & -13\text{dL} \\ 22\text{dL} & 4\text{dL}^2 & 13\text{dL} & -3\text{dL}^2 \\ 54 & 13\text{dL} & 156 & -22\text{dL} \\ -13\text{dL} & -3\text{dL}^2 & -22\text{dL} & 4\text{dL}^2 \end{bmatrix}$
Piezoelectric	Kem_e	$e_{13} b_{\text{p}} \frac{(h_{\text{p}}+h_{\text{b}})}{2}\begin{bmatrix} 0 & -1 & 0 & 1 \end{bmatrix}$
Coupling	Kme_e	Kem_e^T
Case Beam	$EI = E_b I_b$	$\rho hb = \rho_b b_b h_b$
Case Beam+Piezo	$EI = E_b I_b + 2E_p I_p$	$\rho hb = \frac{\rho_b h_b + \rho_p h_p}{h_b + h_p} b_b (h_b + h_p)$

The inertia along the neutral axis are defined for the beam I_b and for the piezoelectric transducer I_p, such as:

$$\begin{cases} I_b = \frac{b_b h_b^3}{12} \\ I_p = \frac{b_p\ h_p^3}{12} + h_p\ b_p \left(\frac{h_b}{2} + \frac{h_p}{2}\right)^2 \end{cases} \tag{3.1}$$

Taking into account N_b elements to assemble elementary matrices with the corresponding boundary condition, the equation of the system from Chapter 2 equation (2.46) is such as :

$$\begin{cases} M\{\ddot{w}\} + J\{\dot{w}\} + K\{w\} + K_{\text{me}}\{V\} = F_{\text{u2}}\{F\} \\ K_{\text{em}}\{w\} - C_p\{V\} = -\{q\} \end{cases} \tag{3.2}$$

Where $M \in \mathbb{R}^{2N_b \times 2N_b}$ is the mass matrix, $K \in \mathbb{R}^{2N_b \times 2N_b}$ is the stiffness matrix, $J \in \mathbb{R}^{2N_b \times 2N_b}$ is the damping matrix, $V \in \mathbb{R}$ is the electric voltage of the piezoelectric transducer, $q \in \mathbb{R}$ is the electric charge, $F_{\text{u}}F \in \mathbb{R}^{2N_b \times 1}$ are the external forces applied on the system, and $w \in \mathbb{R}^{2N_b \times 1}$ are the displacements w_i and θ_i of the system, such as:

$$\{w\} = \begin{bmatrix} w_1\ \theta_{\ 1} \ \dots\ w_{\text{N}_\text{b}}\ \theta_{\ \text{N}_\text{b}} \end{bmatrix}^T \tag{3.3}$$

A state space representation of the system is built, such as:

$$\begin{cases} \dot{X} = \text{AX} + \text{BU} \\ Y = \text{CX} + \text{DU} \end{cases} \tag{3.4}$$

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Where $X = \left\{ \begin{bmatrix} w \\ \dot{w} \end{bmatrix} \right\}$ is the state vector, $U = \left\{ \begin{bmatrix} V \\ F \end{bmatrix} \right\}$ is the input vector, $Y = \left\{ \begin{bmatrix} q \\ w \end{bmatrix} \right\}$ is the output vector and matrices A, B, C and D are defined, such as:

$$A = \begin{bmatrix} 0 & \mathrm{Id} \\ -M^{-1}K_{\mathrm{eq}} & -M^{-1}J \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 \\ -M^{-1}K_{me} & M^{-1}F_{\mathrm{u}} \end{bmatrix}$$
$$C = \begin{bmatrix} -K_{\mathrm{em}} & 0 \\ \mathrm{Id} & 0 \end{bmatrix} \quad D = \begin{bmatrix} C_p & 0 \\ 0 & 0 \end{bmatrix}$$

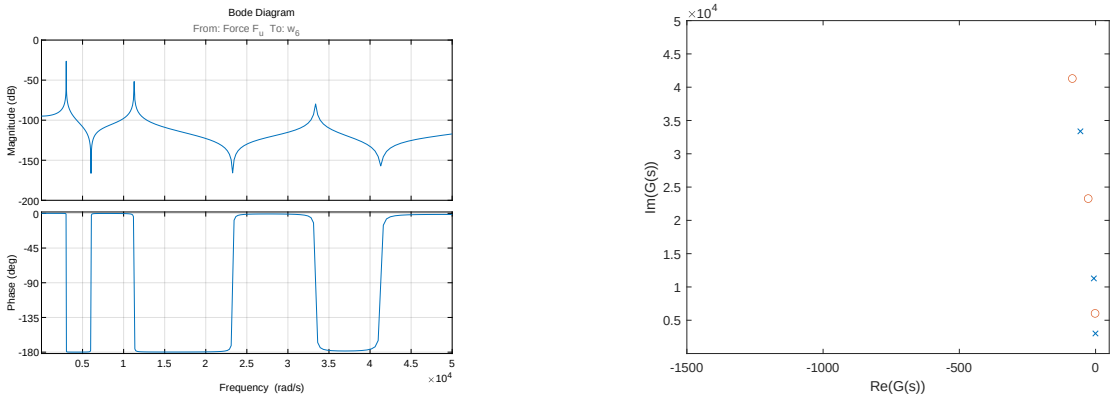
The system can also be written with a transfer function representation $G(s)$ in the Laplace domain, such as:

$$G_{tot}(s) = C(sId - A)^{-1}B + D \tag{3.5}$$

Where s is the Laplace variable.

3.1.2 System response and disturbance

The model is built with $N_b = 10$ elements, including five beam piezo elements located from point 2 to point 6. It is chosen here to focus on the first three modes of the system. Figure 3.2a shows the bode diagram of the system from F_6 to w_6, taken as the reference input-output actuator-sensor. The system is slightly damped thanks to the matrix $J = 1e{-}07K$, so that the real part of the poles and zeros is not equal to zero (Figure 3.2b).



(a) (b)

Figure 3.2: System $G(s)$ slightly damped: Bode diagram (a) and pole (blue ×) / zero(red o) pattern (b) of the structure with collocated actuator and sensor with $N_b = 10$.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

3.1.3 Passive Control implementation

The objective is to apply passive methods presented in the bibliography to damp the system and reduce the system vibrations at the selected frequencies. An impedance circuit is applied to the transducers electrodes acting as a feedback loop controller $C(s)$ on the system $G(s)$ as introduced in Figure 3.3.

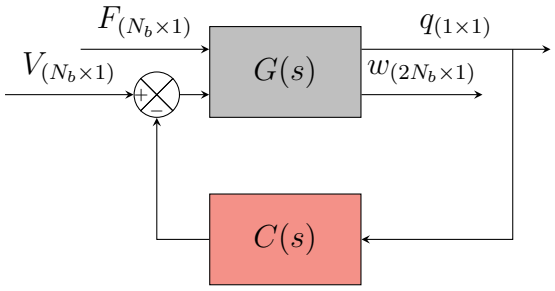


Figure 3.3: Control loop block diagram

The impedance, which is a shunt resistor, is input defined by

$$C(s) = R_{shunt} s \tag{3.6}$$

Where s is the Laplace variable and R_{shunt} is the impedance set between the two transducers electrodes. The system is disturbed with a white noise on input F_2. The system response can be measured by the other collocated transducer q or with the displacement w.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

3.2 Proposed Methodology applied on the example case

This thesis proposes a methodology using reinforcement learning algorithm to tune a predefined control law on a smart structure. The main steps are developed on the example case with the definitions of the components as summarized in Figure 3.4. An episode of the training occurs as follows :

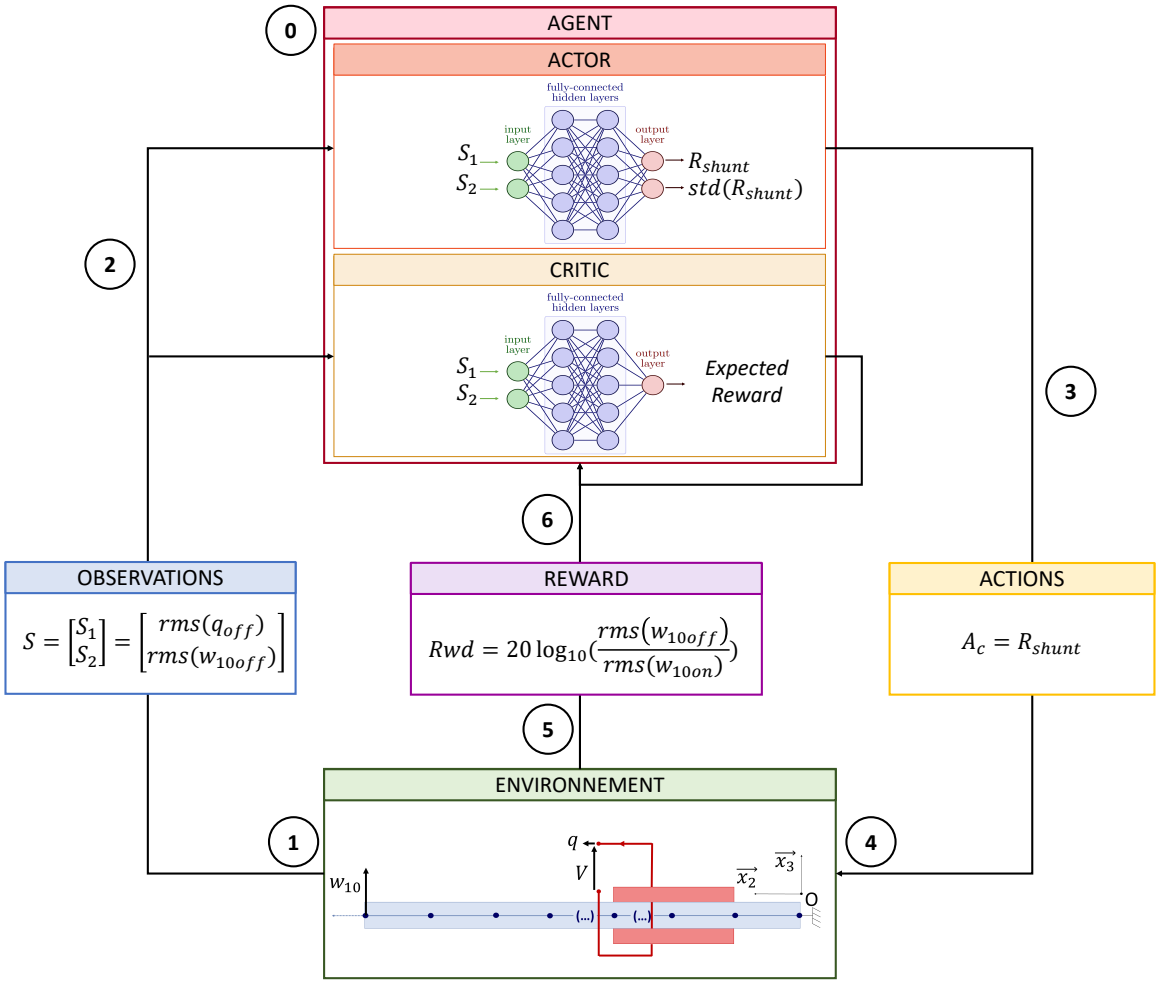


Figure 3.4: Reinforcement Learning Algorithm applied on the 1D smart beam

- (0)-**Agent initialization**: The actor and the critic neural networks are initialized with random weight and bias.
- (1)-**Environment initialization**: The beam response to a white noise is simulated with the short circuit condition $V = 0$ as the control off state.
- (2)-**Observation Computation**: The observations are calculated using chosen functions to convert the time domain data to representative values of the simulated environment.
- (3)-**Agent decision process**: The agent estimates the expected reward with the critic and an action value according with the input observations.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

- (4)-**Environment running**: The beam response to a white noise is simulated with the condition $C(s) = sR_{shunt}$.
- (5)-**Reward computation**: The defined reward function is calculated using the time domain values from step (1) and step (4) simulations.
- (6)-**Agent update**: The actor and critic weights and biases are updated using TRPO or PPO algorithms (Chapter 2 part 2.2.2), taking into account the expected reward from step (3) and the reward from step (6).

The definitions of each element are detailed in the following lines.

3.2.1 Environment Definition

As seen in Chapter 2 part 2.2.2.1, the reinforcement learning algorithm does not use any database, but needs an environment to process trial errors and train an agent. For the example case, the environment is defined by a 1-dimensional finite element cantilever beam (part 3.1). To communicate with the agent, observations, actions, and rewards must be defined.

3.2.1.1 Observations

The observations must be representative of the structure and contain significant information. The use of time-varying data is not relevant in the vibration context because the goal is to minimize the overall response. Well-known metrics that are representative of the time signal can be used. Here, it is decided to use the root mean square of the finite element model output: q the current at the transducer electrodes and w_{10} the displacement at the free end of the beam, such as:

$$S = \begin{cases} S_1 = rms(q) \\ S_2 = rms(w_{10}) \end{cases} \tag{3.7}$$

3.2.1.2 Action

The actions must be defined according to the control law to be applied to the structure. For this simple application, it is chosen to apply a resistive shunt to the piezoelectric transducers. The resistive shunt range values R_{shunt} are defined as a continuous space between set minimum R_{min} and maximum R_{max} values, such as:

$$R_{shunt} \in \{R_{min} = 10^0; R_{max} = 10^{12}\}[\Omega] \tag{3.8}$$

To ensure that the output of the neural network is of the same order of magnitude, the actions are quantified by powers of 10, so that the output of the neural network is between 0 and 12.

3.2.1.3 Reward

The reward function must be designed in accordance with the control objective. With the use of a resistive shunt, the goal is to minimize the vibration of the beam. The bibliography (Chapter 2 Part 2.1.2.2), shows that an optimal value can be reached when the system displays maximum damping for the 1st mode, so when the time response of

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

the beam at the free end is minimum. The chosen reward $Rwd \in \mathbb{R}$ to maximize is defined to have a comparison between the free and controlled oscillations of the beam at the free end, such as:

$$Rwd = 20log_{10}(\frac{rms(w_{10off})}{rms(w_{10on})}) \tag{3.9}$$

Where w_{10off} is the time response of the beam at the free end when the control is off and w_{10on} when the control is on. This definition allows the use of complex signals, not just harmonics. Since the reward varies according to one variable, the resistance value, it is possible to have a graphical representation of it on Figure 3.5. This function form is consistent with the expected behaviour:

- (1) If the resistance value is small, it can be assimilated to a short-circuit condition.
- (2) The first mode is then damped proportionally to the increase in resistance until an optimum is reached.
- (3) The resistance is high enough to be assimilated to an open circuit electrical configuration.

Note that the reward variation is convex according to the resistance value and that the problem could easily be solved with the classical optimization algorithm seen in Chapter 2 part 2.1.3, but for the example a reinforcement learning algorithm is implemented.

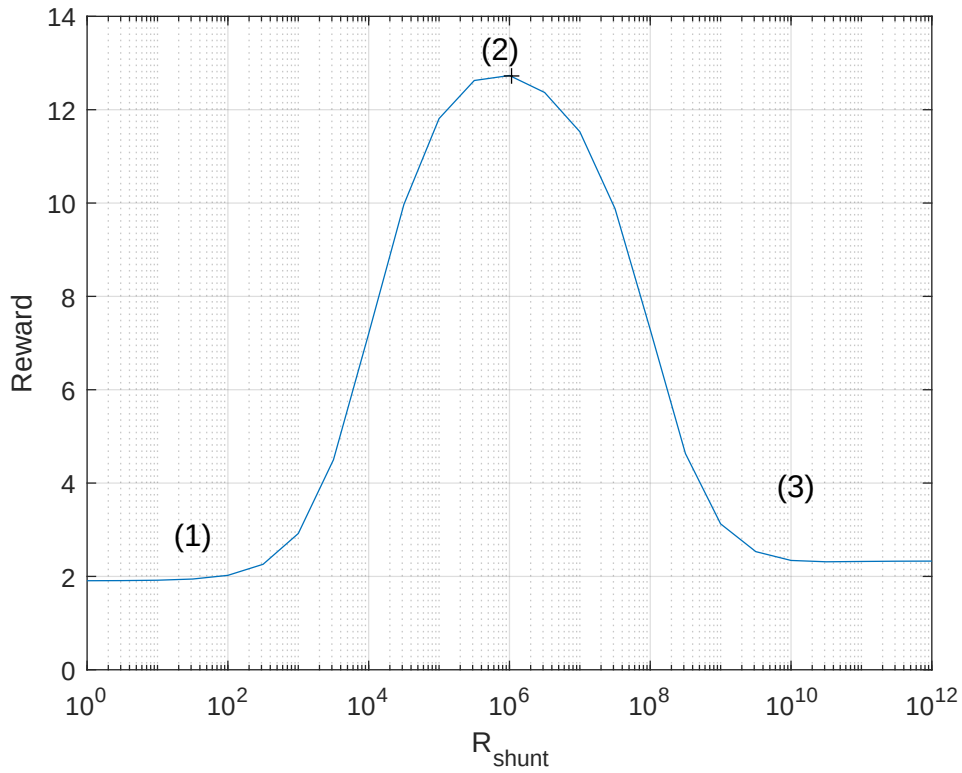


Figure 3.5: Reward: (blue line) variation with resistance and (black "+") resistance tuning using Hagood (Chapter 2 Part 2.1.2.2).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

3.2.2 Reinforcement Learning Algorithm

As seen in Chapter 2 part 2.2.1.2, there are many reinforcement algorithms. Since in this example the observation values and action values are continuous, the TRPO and PPO methods are used to train the agent composed of a critic and an actor.

3.2.2.1 Hyperparameters

Every reinforcement learning method has a set of tuning parameters called hyperparameters. Usually default parameters are efficient enough, but tuning them will influence the learning efficiency of the actor and critic neural networks. Table 3.3 summarizes the most important hyperparameter values and their effect on training.

Table 3.3: Hyperparameter definition for PPO and TRPO algorithms

Hyperparameter	Name	Default value	Range values	Effect
Enthropy Loss Weigth	$w \in \mathbb{R}$	0.01	$\{0; 1\}$	It enhances agent exploration with high entropy loss.
Learning rate	$\alpha \in \mathbb{R}$	0.01	$\{0; \infty\}$	Training may lead to sub-optimal results or divergence if the learning rate is too high. If too low, training will take a long time to complete.
Discount Factor	$\gamma \in \mathbb{R}$	0.99	$\{0; 1\}$	In the learning process, it reduces the influence of the expected reward calculated by the critic.
Clip Factor	$\epsilon \in \mathbb{R}$	0.2	$\{0; 1\}$	Used only in the PPO, it limits the amount of change in each step of the policy (actor NN) update.

3.2.2.2 Episode Definition

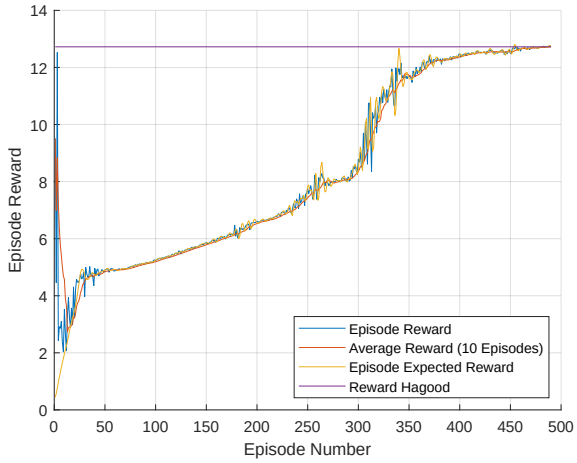
The training process is divided into episodes. An episode can be divided into many steps. For example, if the algorithm learns to play Chess, an episode will be a full game until one of the players loses, and a step will be moving a piece. In this example, an episode consists of only one step because the environment does not change. As seen on Figure 3.4, an episode always starts with an initialization of the environment to inform the algorithm about the initial state S. In this example, it corresponds to the structure response to a white noise input at F_2. According to the state values, the agent will estimate an action A_c to perform on the environment. According with the environment response, a reward value is returned to the agent to update its NN parameters.

3.2.2.3 Training

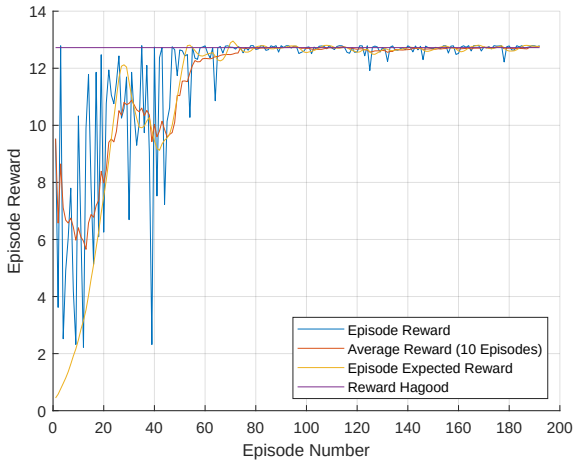
Episodes are performed with the goal of maximizing the reward. Training stops when either the maximum defined reward or the maximum defined number of episodes is reached. For the example case, the maximum reward is calculated for the optimal resistance value defined by Hagood (Chapter 2 Part 2.1.2.2). The maximum number of episodes is arbitrarily set to 1000. Figure 3.6 shows the training of PPO and TRPO algorithms. Both

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

algorithms succeed in reaching the maximum fixed reward to end the training. As expected, TRPO needs more episodes because the learning is more progressive than PPO, which covers a larger range of values at the beginning of the training.



(a)



(b)

Figure 3.6: Training: (a) TRPO, (b) PPO.

3.3 Control Performances

To quantify the performance of the shunt tuned with the Deep Reinforcement Learning (DRL) algorithm, the system response is shown in Figure 3.7. Figures 3.7a and 3.7b show a difference in damping efficiency on the first mode between the two DRL methods: TRPO tuning introduces less attenuation than PPO tuning. However, this difference is not significant in the time domain when the system is disturbed by the input noise (Figure 3.7c). This is explained by the reward definition, which uses time-domain metrics and thus includes all frequencies excited by the noise perturbation. Finally, the PPO and TRPO algorithms are capable of tuning a resistance value that is as efficient as Hagood and Preumont methods.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

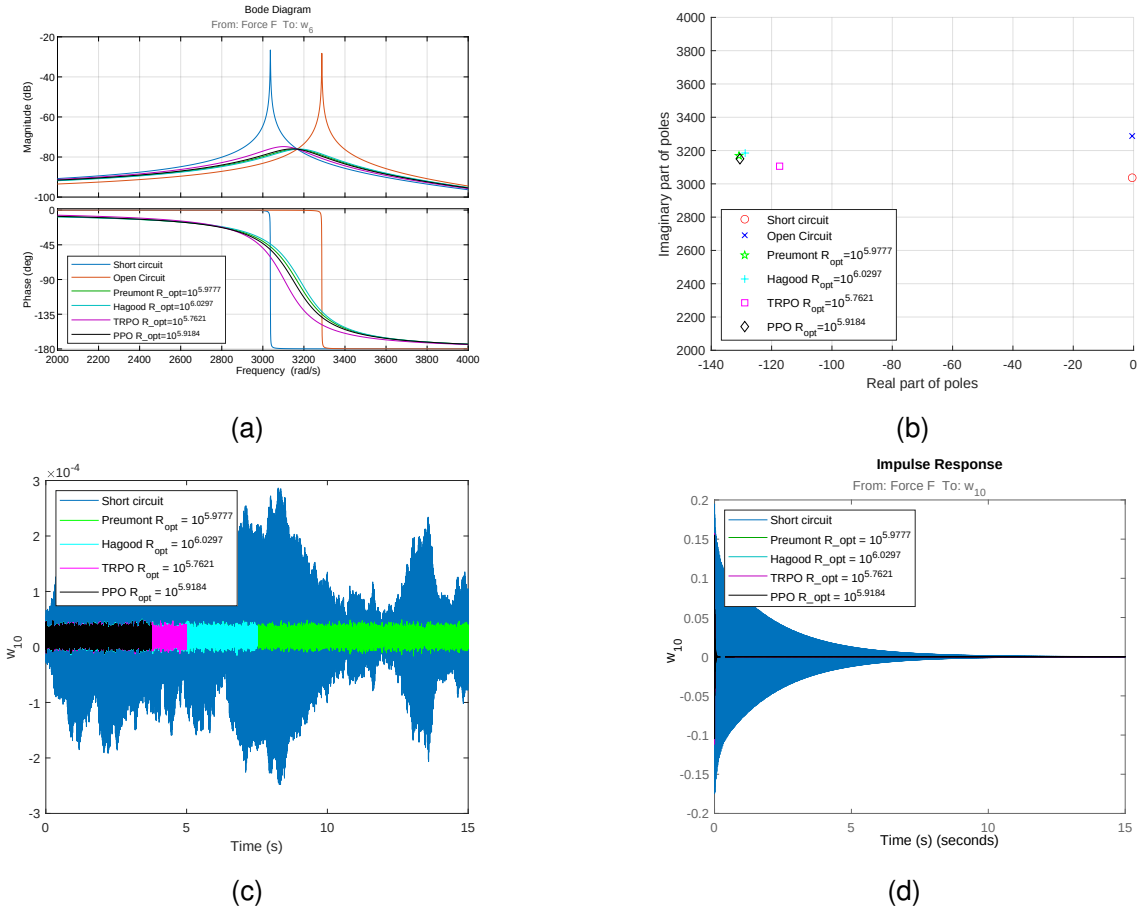


Figure 3.7: Control Performances on the example system: (a) Bode response of G(s), (b) Complex plan, (c) Impulse and (d) Noise response.

3.4 Conclusion

The use of reinforcement learning tools in a numerical environment has proven to be efficient on many application cases in the bibliography [123]-[68]. Applied to the example case, it is able to tune a resistive shunt similar to the one obtained following the bibliography. The challenges are known to test this methodology in more complex situations. First it is implemented on an experimental setup with only one actuator (Chapter 4 Part 3). Then, it is tested with the tuning of several variables (Chapter 4 Part 3.4.2). Once the efficiency of the method is proven in SISO, it is applied to MIMO systems to tune ether active (Chapter 5) or passive control (Chapter 6). In these cases, there are no analytic formulations to obtain optimal control. The goal of using artificial intelligence in this case is to tune efficient distributive controllers according to the mitigation objectives. The challenges and limitations of the method are highlighted at each step of the process and summarized at the end of the document in the conclusion part (Chapter 7).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Chapter 4

Tuning active control laws on a smart cantilever beam: experimental study

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

This chapter proposes the active control of a smart cantilever beam composed of two collocated piezoelectric transducers.The methodology presented in the previous chapter is applied here in the context of an experimental test bench. The system under study is modeled by a 3-dimensions finite element model and compared with an experimental setup. The control laws are selected and tuned according to the bibliography exposed in Chapter 2 to mitigate the free-end vibrations of the beam. It is decided to implement a derivative controller, considering one parameter to be tuned and then a lead controller with three parameters. A Deep Reinforcement Learning (DRL) methodology is presented and applied to the system to tune the control laws. A comparison between the different approaches in terms of control performance is developed in the results part. The advantages and limitations of DRL in terms of robustness to structure and input signal variations are discussed.

4.1 Definition and design of a smart cantilever beam with two collocated piezoelectric patches

Before any control is applied to a structure, it must be designed to optimize the effect of the transducers on the desired modes. A finite element model is created using COMSOL software to determine the beam dimensions so that the coupling coefficient with the piezoelectric transducers is improved.

4.1.1 Finite element model

4.1.1.1 Model definition

The 3-dimensions finite element model is used to model the frequency response of an aluminum cantilever beam composed of two piezoelectric transducers.

Variables and material properties

The material properties and dimensions of the transducers are given by the manufacturer. The dimensions to be adjusted are the length L_b, width b_b and thickness h_b of the aluminum part and the position l_d of the piezoelectric element on it. The main dimensions and material properties are given in Table 4.1 and refer to the variables described in figure 4.1. The beam dimensions are unknown and must be tuned to optimize the piezoelectric coupling coefficient.

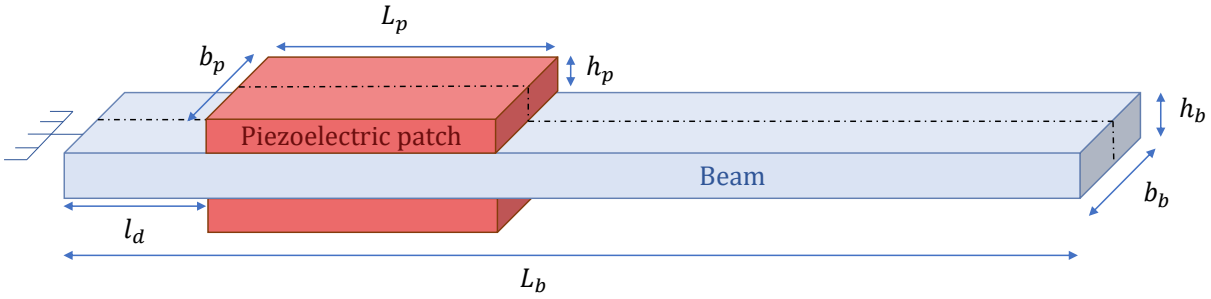


Figure 4.1: Scheme: cantilever beam dimensions

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Table 4.1: Constant definitions

	Variable	Beam		Piezoelectric transducer (PZ26)		Unit
		Name	Value	Name	Value	t
Dimension	Length	L_b	to optimize	L_p	0.05	$[m]$
	Width	b_b	to optimize	b_p	0.05	$[m]$
	Thickness	h_b	to optimize	h_p	0.001	$[m]$
	Location	-		l_d	to optimize	$[m]$
Material properties	Young modulus	E_b	70e9	c_{11}^S	1.68e11	$[Pa]$
				c_{12}^S	1.10e11	$[Pa]$
				c_{13}^S	9.99e10	$[Pa]$
				c_{33}^S	1.23e11	$[Pa]$
				$c_{44}^S = c_{55}^S$	3.01e10	$[Pa]$
				c_{66}^S	2.88e10	$[Pa]$
	Density	ρ_b	2700	ρ_p	7700	$[kg.m^{-3}]$
	Coupling constant	-		e_{31}	-2.8	$[C.m^{-2}]$
		-		e_{33}	14.7	$[C.m^{-2}]$
		-		e_{15}	9.86	$[C.m^{-2}]$
	Relative permittivity	-		ϵ_{11}^S	828	$[-]$
		-		ϵ_{33}^S	700	$[-]$

Initial and boundary conditions

All components are defined in the linear elastic domain according to the piezoelectric equations defined in Chapter 2, Part 2.1.1.2. All initial displacements are set to zero and only one face of the beam is constrained by an embedding condition. In the piezoelectric elements, the electric charge is conservative and the initial stress is set to zero. The electrical mass is placed on the faces close to the beam and the electrical potential is integrated on the electrode of the other face to match the equipotential condition (Chapter 2 part 1.1.2 equation 35).

Mesh

Different default mesh sizes are proposed by the software (COMSOL), varying from large to small elements. Figure 4.2 shows an example of four of them applied to the structure. A mesh analysis is performed to see its influence on the estimation of the first frequency mode with short-circuit and open-circuit configuration (figure 4.3). As expected, the computation time increases with the number of elements. The frequency varies linearly compared to the exponential variation of the computational time with the mesh size. It should be noted that anti-resonances are in fact even more sensitive to mesh size, but this is not investigated here. The mesh size has a small influence on the coupling coefficient k, which varies from 0.24 for the extremely coarse mesh to 0.22 for the finer mesh. Finally, the best compromise between computation time and accuracy, is the normal mesh size with 20000 elements. It guarantees an absolute error of less than 1% in frequency compared to the finest mesh in just 7 seconds of simulation.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

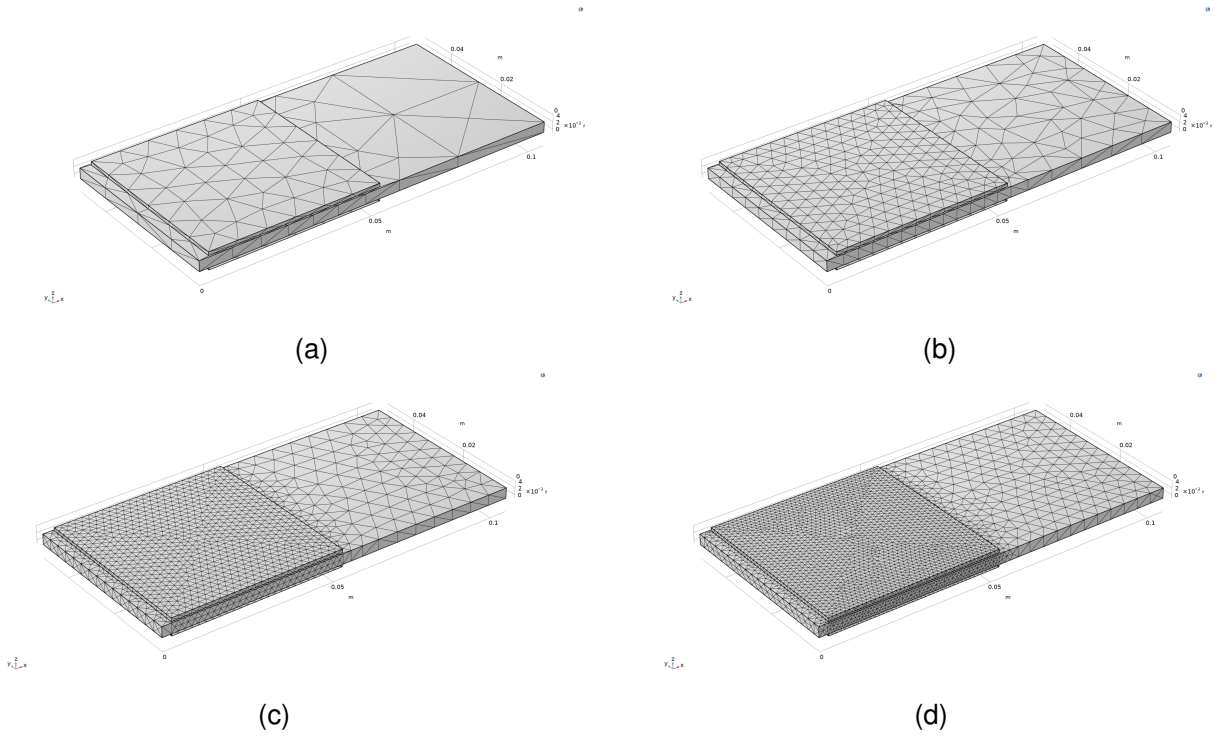


Figure 4.2: Finite element model of the cantilever beam with different meshes: (a) extremely coarse, (b) coarser, (c) normal, (d) finer.

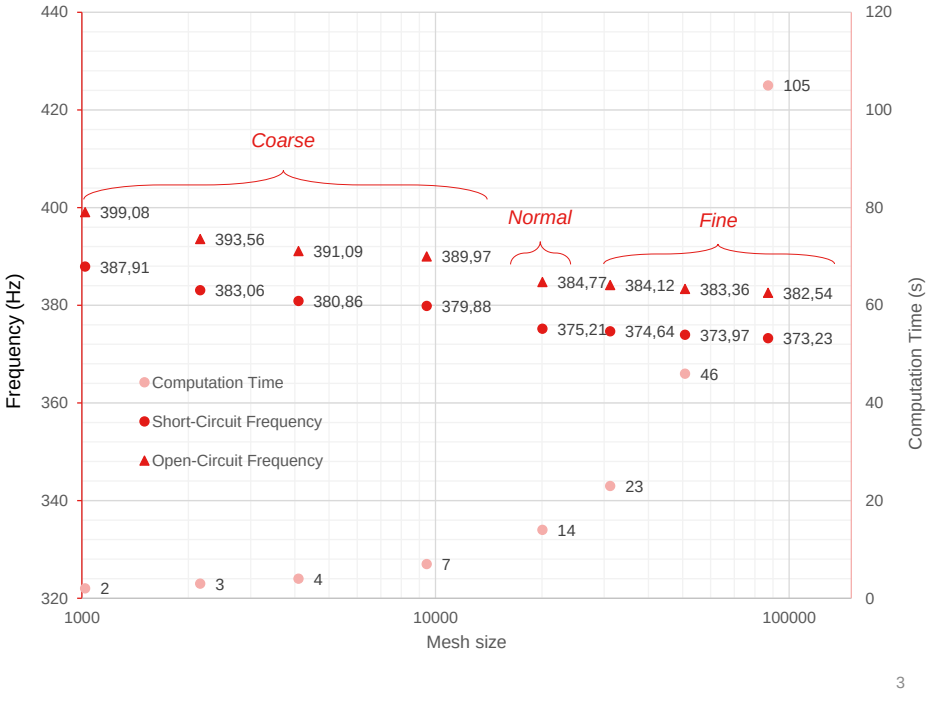


Figure 4.3: Influence of the mesh on the short and open circuit frequencies (left axis) and computation time (right axis).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

4.1.1.2 Optimization and modal study

Optimize coupling coefficient

A parametric approach is used to find the beam dimensions that maximize the piezoelectric coupling coefficient (Chapter 2, Part 1, Equation 42) for the first bending mode. Figure 4.4 shows the variation of the coupling coefficient as a function of the beam dimensions in the three spatial directions. The optimal ratio is close to 0.6 for the length, 0.9 for the width and 0.3 for the thickness. According to figure 4.4d, it is possible to slightly improve the coupling coefficient by moving the position slightly away from the embedding end of the beam.

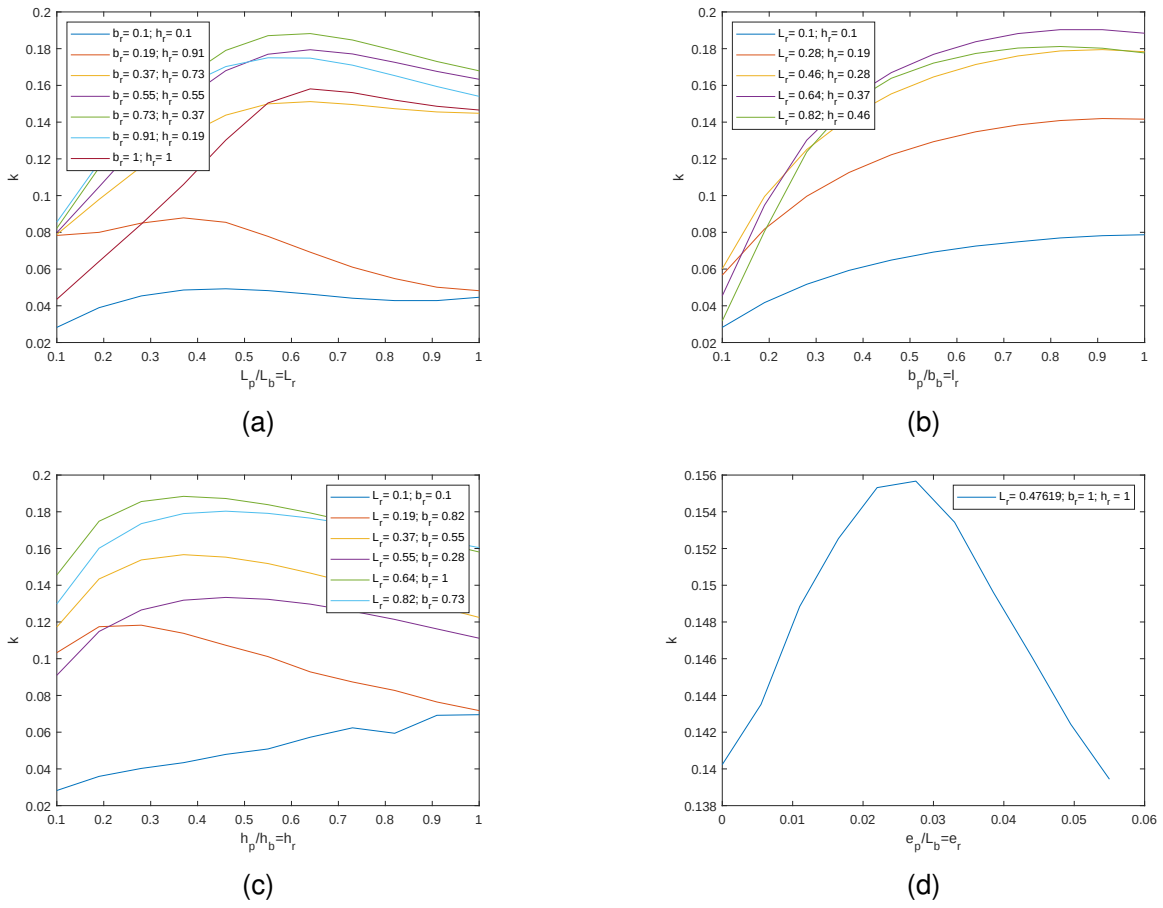


Figure 4.4: Coupling coefficient variations with beam dimensions: (a) length, (b) width, (c) thickness ratio where $e_p = 0$ and (d) transducers position.

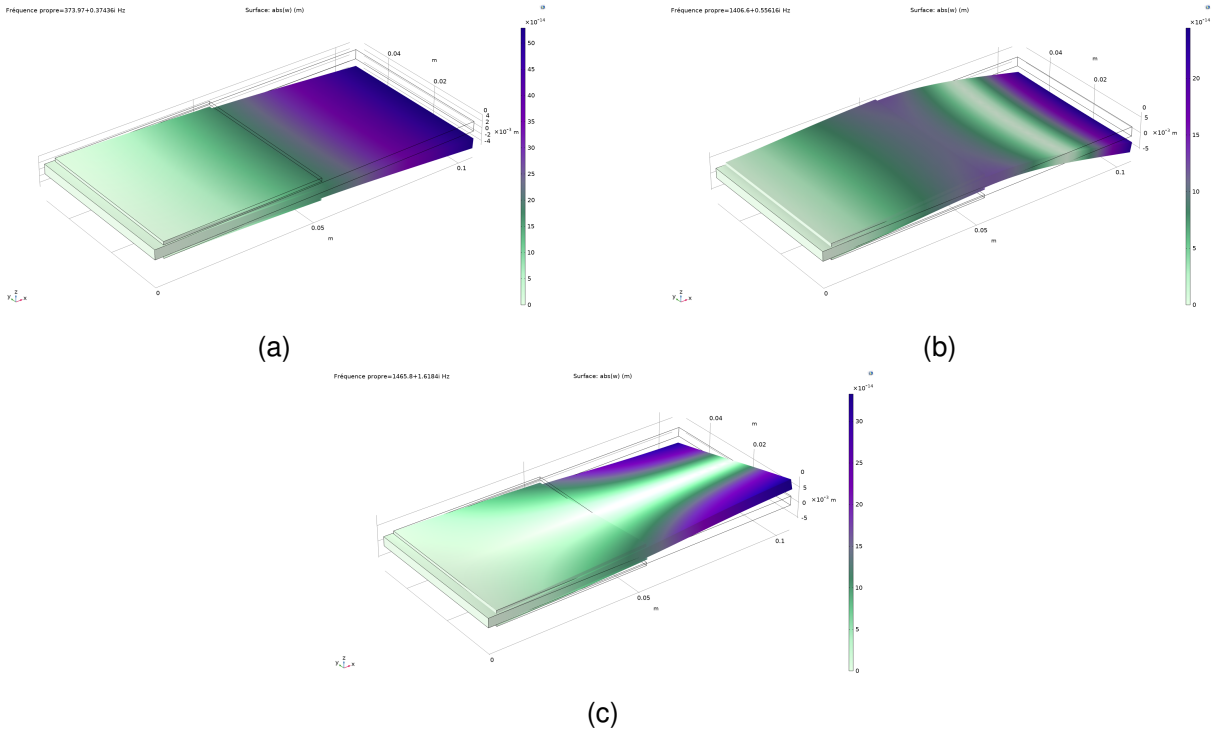
Modal response

The dimensions are set as close to the optimum as possible to perform a modal analysis of the beam (Table 4.2) with an effective expected coupling coefficient $k = 0.155$. The first three modes are identified in the figure 4.5. The first two modes are bending modes and the third is a torsional mode. According to the physical properties and modal shapes, the transducers have no effect on the third mode.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Table 4.2: Setup dimensions

	Dimensions			Materials
	Name	Value	Unit	
Beam	L_b	105	[mm]	Aluminium
	b_b	52	[mm]	
	h_b	3	[mm]	
Piezoelectric Patch	L_p	50	[mm]	Pz26
	b_p	50	[mm]	
	h_p	1	[mm]	
	e_d	3	[mm]	



(a) (b)

(c)

Figure 4.5: Mode shape visualization: (a) mode 1 (first bending mode), (b) mode 2 (second bending mode), (c) mode 3 (first torsion mode).

4.1.2 Experimental validation

The smart structure is built to compare the FEM analysis with the experimental properties.

4.1.2.1 Experimental setup

Bonding process

The Smart Beam fabrication uses a bonding process that is divided into key steps as shown in figure 4.6.

- Figure 4.6a: The surface is sanded to introduce roughness to the aluminum surface.
- Figure 4.6b: All sanding residues are cleaned with alcohol to prevent dust from melding with the adhesive.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

- Figure 4.6c: The sanded surface is coated with an epoxy resin before the transducer is attached to it with a sliding motion to avoid air bubbles in the thin layer of adhesive.
- Figure 4.6d: The structure is then placed in a vacuum bag and placed under a heat lamp to allow the resin to dry until it is completely polymerized.

During the drying phase, it is essential to keep the transducer immobile on the beam to avoid slippage.
Since the bonding process is done by hand, it introduces uncertainties mainly in the use of the resin. This element is not taken into account in the FEM model and can introduce slight damping to the structure and affect the piezoelectric coupling coefficient. Also, this process does not guarantee a constant thickness of resin over the entire surface of the transducer. Contact between the electrode and the aluminum beam may occur, resulting in an electrical short. In order to avoid this, the electrical mass must be defined on the electrodes facing the beam and be in contact with the beam.

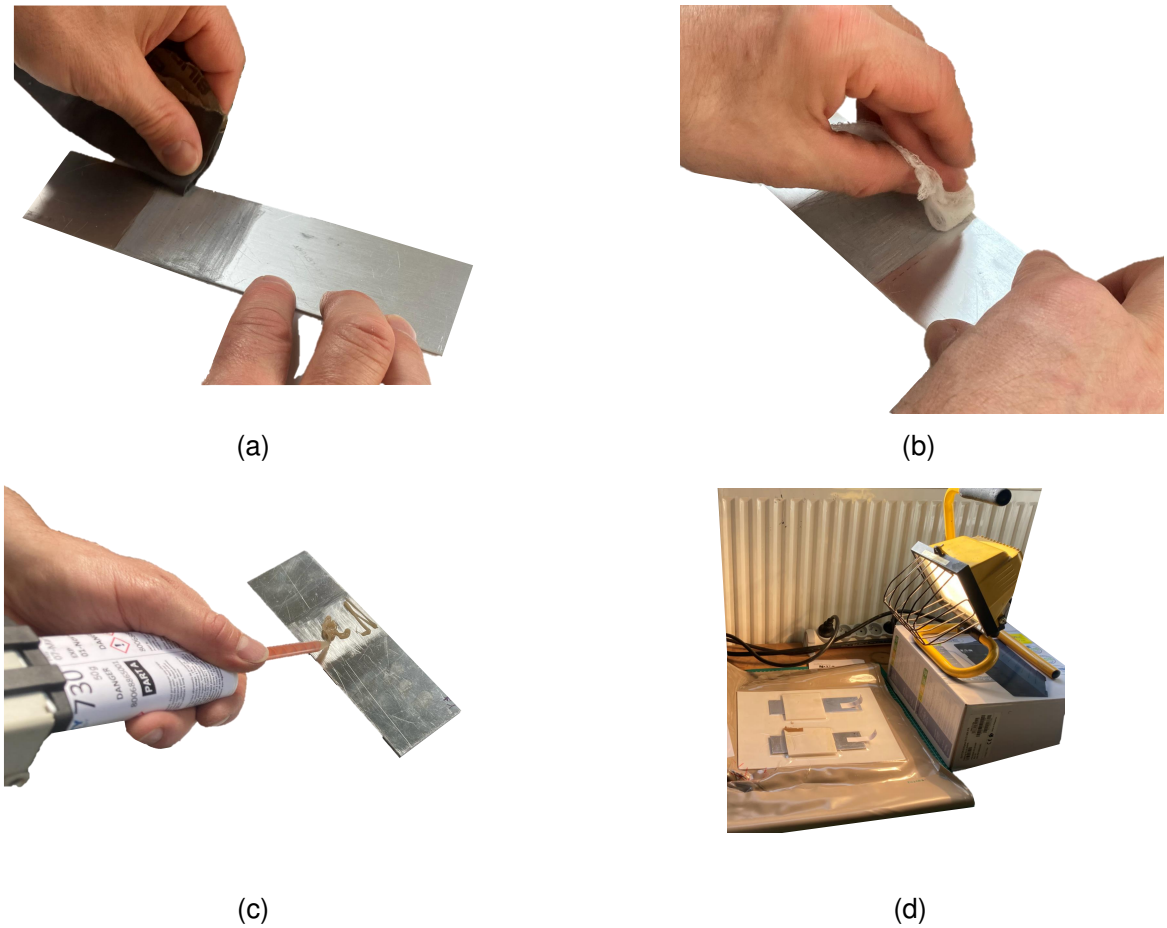


Figure 4.6: Key steps for sticking the transducer to the beam: (a) surface sanding , (b) alcohol cleaning, (c) applying and spreading the adhesive and, (d) vacuum drying.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Setup presentation

The experimental setup is composed of the designed smart beam with two collocated piezoelectric patches close to the clamped end as displayed in figure 4.7. The beam has length L_b, width b_b, and thickness h_b. The two collocated piezoelectric elements are attached to the beam at a distance e_d from the supported end. The piezoelectric transducers dimensions are length L_p, width b_p, and thickness h_p. All dimensions are already given in Table 4.2 in part 4.1.1.1.

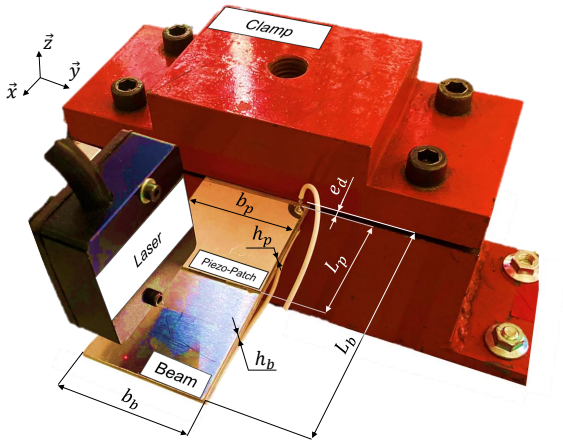


Figure 4.7: Experimental setup: beam

One piezoelectric transducer is set as an actuator with v as the input signal and the other as a sensor monitoring an image of the beam bending curvature x_1. A laser is used to monitor the displacement of the free beam end x_2. All devices used to run the experiment are described in figure 4.8 and the hardware information is given in Table 4.3.

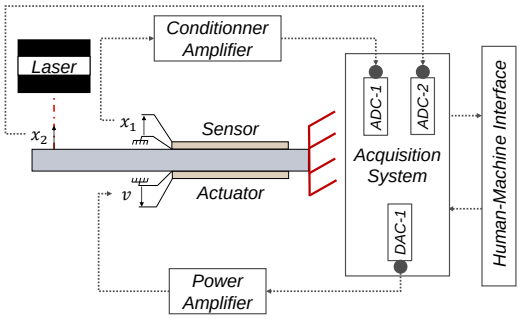


Figure 4.8: Experimental setup: data acquisition systems

4.1.2.2 Experimental results

Coupling coefficient

The experimental setup introduces new uncertainties due to the clamping element and bonding layers that are not accounted for in the FEM model. Experimental modal identifications are performed to compute the experimental coupling coefficient with the method introduced in Chapter 2 Part 1.2.2. Two configurations are used because the structure is composed of two patches:

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Table 4.3: Experimental devices

Device	Brand	Reference	Sensitivity	Measuring Range
Laser	Micro-Epsilon	LD1607-2	10V/mm	2mm
Acquisition System	DSPACE	DS1104	-	+/-10V
HMI	DSPACE	ControlDesk 7.6	-	-
Conditionner Amplifier	BK	Type 2626	1pC/V	-
Power Amplifier	PI	DuraAct Piezo Amplifier	$<$10 mV	Input Voltage : -4 to 10 V Outpur Voltage : -100 to 250 V

- Patch A and B: The coupling coefficient is calculated from the frequencies identified when the two patches are open and short-circuited at the same time. The structure is disturbed with a repetitive impulse induced by a hammer.
- Patch A or B: The coupling coefficient is calculated from the frequencies identified when only one patch is open and shorted. During experiment, the other patch is shorted or used to induced white noise perturbation to the structure.

The results are shown in the following Table 4.4 and summarize the mean coupling coefficient with uncertainties calculated from the different measurements and analysis methods.

Table 4.4: Coupling coefficients on the first mode

Study	Elements	Electrodes	Frequency (Hz)	Coupling Coefficient (k)
FEM	Patch A and B	Short-Circuit	375.21	
	Patch A and B	Open-Circuit	384.77	0.23
	Patch A or B	Open-Circuit	380.14	0.16
Measurement	Patch A and B	Short-Circuit	240.08 ± 0.45	
	Patch A and B	Open-Circuit	241.06 ± 0.84	0.10 ± 0.01
	Patch A	Open-Circuit	240.80 ± 0.25	0.09 ± 0.03
	Patch B	Open-Circuit	241.21 ± 0.26	0.10 ± 0.03

The experimental methods can only identify the coupling coefficient of the first mode. When the hammer is used, the higher frequency modes are masked with background noise due to the laser sensitivity. The experimental standard deviations overlap because the various estimation methods using vibration measurements introduce uncertainties depending on the signal processing used. Also, the open and short frequencies are close to each other (less than 1 Hz) and their identification is closely related to signal processing and analysis. Considering the uncertainties, the experimental coupling coefficient is close to 10 percent for the first mode at 240 Hz. It is two times lower than the expected one obtained with FEM. In contrast to the modeled results, the effect of two patches was no better than that of one. The measurement protocol should be reviewed to provide more accurate results.

The estimated frequencies are also 120 Hz higher than the measured ones. These differences partially explained by several factors, such as the experimental construction process, the clamping position and temperature variations. During the experiments, a

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

variation of more than 40 Hz was observed between two measurement periods carried out in variation of media humidity and temperature conditions. A parametric study should be carried out to identify the influence of each physical parameter on the modeling, but this is beyond the scope of this thesis.

The differences between the FEM and the experimental setup cannot be neglected in the context of tuning an efficient controller to mitigate vibrations. Deep reinforcement learning algorithms need a reliable representation of the system in the environment definition, so another strategy is investigated to model the structure.

System identification

Instead of using an FEM model, which is not very close to reality, it is decided to use an experimental model based on measurements. To do so, a band-limited noise $v(t)$ defined in equation (4.1) is used as an input for the identification process and is applied to the actuator transducer with a noise power equal to 0.1, a sampling frequency $F_s = 10$ kHz, and a second order low pass filter considering a cut-off frequency $F_c = 1000$ Hz:

$$v_{d1}(t) = Rnd(t) * \mathcal{L}^{-1}\{\left(\frac{2\pi F_c}{s + 2\pi F_c}\right)^2\} \tag{4.1}$$

Where $\mathcal{L}^{-1}$ is the inverse Laplace transform and $*$ is the time-domain convolution. The measured signals x_1 and x_2 allow the identification of two Single Input Single Output (SISO) systems G_1 and G_2 defined by the following transfer functions:

$$\begin{bmatrix} X_1(s) \\ X_2(s) \end{bmatrix} = \begin{bmatrix} G_1(s) \\ G_2(s) \end{bmatrix} V(s) \tag{4.2}$$

Where $X_1(s)$, $X_2(s)$ and $V(s)$ are the Laplace transform of x_1, x_2 and v respectively.

To build the experimental model from the measured frequency response functions, poles and zeros estimation is chosen, such as:

$$G_{i=\{1;2\}}(s) = \frac{\kappa}{s^{n_0}} \frac{\Pi_{i_z=1}^{n_z}(1 - z_{i_z}^{-1}s)}{\Pi_{i_p=1}^{n_p}(1 - p_{i_p}^{-1}s)} \tag{4.3}$$

Where $\kappa \in \mathbb{R}$ is a constant gain, $n_0 \in \mathbb{Z}$ is the number of null poles (integrator if $n_0 > 0$, derivative if $n_0 < 0$), $z_i \in \mathbb{C}$ zeros and $p_i \in \mathbb{C}$ poles of the system, $n_z \in \mathbb{N}$ the number of zeros and $n_p \in \mathbb{N}$ the number of poles, G_i must be proper ($n_p + n_0 \geq n_z$).

The numerical transfer function for G_1 has 12 poles and 10 zeros, while that for G_2 has only 7 poles and 6 zeros, because higher frequencies are not measured, in a frequency range from 10 to 5000 Hz. The pole and zero values set for each transfer function are given in the Appendix B.

It is important to consider the spillover effect for higher frequencies, even if the input disturbance is below 1000 Hz. The transfer functions obtained from experimental measurements and the identified model are shown in Figures 4.9a and 4.9b. Within this frequency range, the piezoelectric sensor can observe four modes, but the laser can only detect two.

The spectral coherence between the two piezoelectric transducers, as shown in Figure 4.9c, is close to one within the desired frequency bandwidth. Between the piezoelectric

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

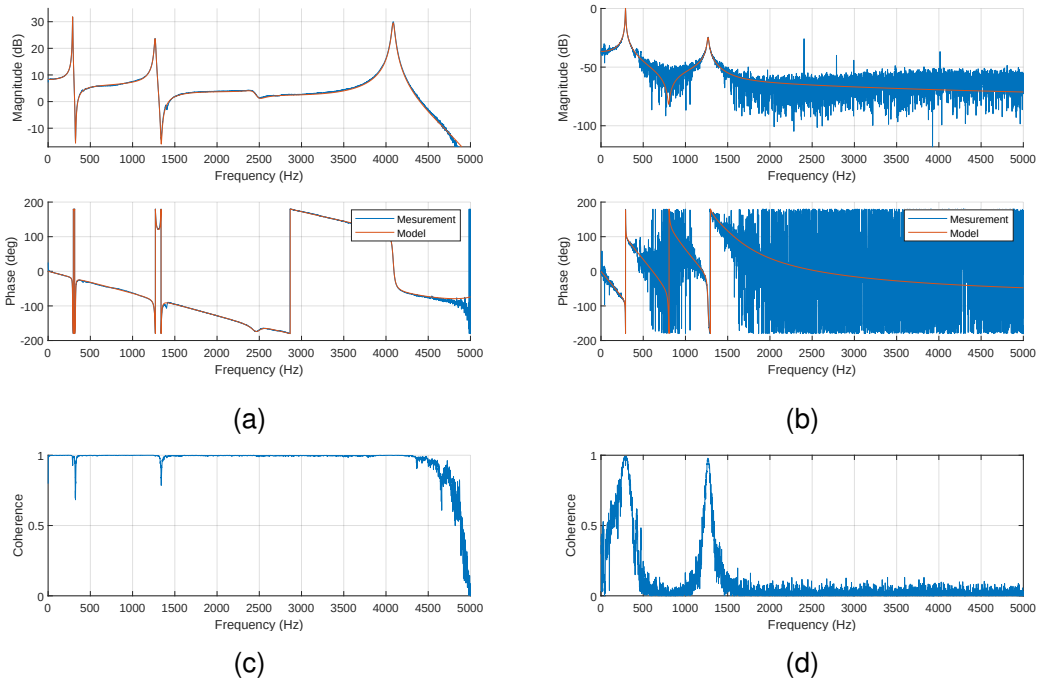


Figure 4.9: Bode diagram with transfer function estimated from measurements (blue line) and the identified model (red line) between 0-5000 $[Hz]$: (4.9a) G_1 and (4.9b) G_2 with (4.9c-4.9d) their corresponding spectral coherence

actuator and the laser, the spectral coherence shown in Figure 4.9d is high only at frequencies close to the natural frequencies of the beam. The signal is very noisy due to the low sensitivity of the laser sensor.

A pair of zeros is added at 660 Hz to fit the experimental case even if it cannot be clearly identified on the measured frequency response function due to noise and laser sensitivity. To reduce this noise, the input power of the disturbance has been increased until saturation due to the acquisition system voltage limitations.

A linear drop in phase is observed and can be explained by the delays introduced by the various elements of the measurement chain, in particular the conditioners and amplifiers of the piezoelectric transducers.

To ensure the model fits the experimental setup, the output signals of the model and the experiment need to be as close as possible. Figure 4.10 displays the time signals from the sensors in the case of band-limited white noise disturbance, with identical inputs for both the experiment and the identified system.

The normalized RMS error between the model output and the experimental structure is 0.11 for x_1 and 0.28 for x_2. This difference is due to measurements and transfer function estimations but it is negligible for the purpose of this work. Therefore, it can be concluded that the model response tendency fits the experimental behavior.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

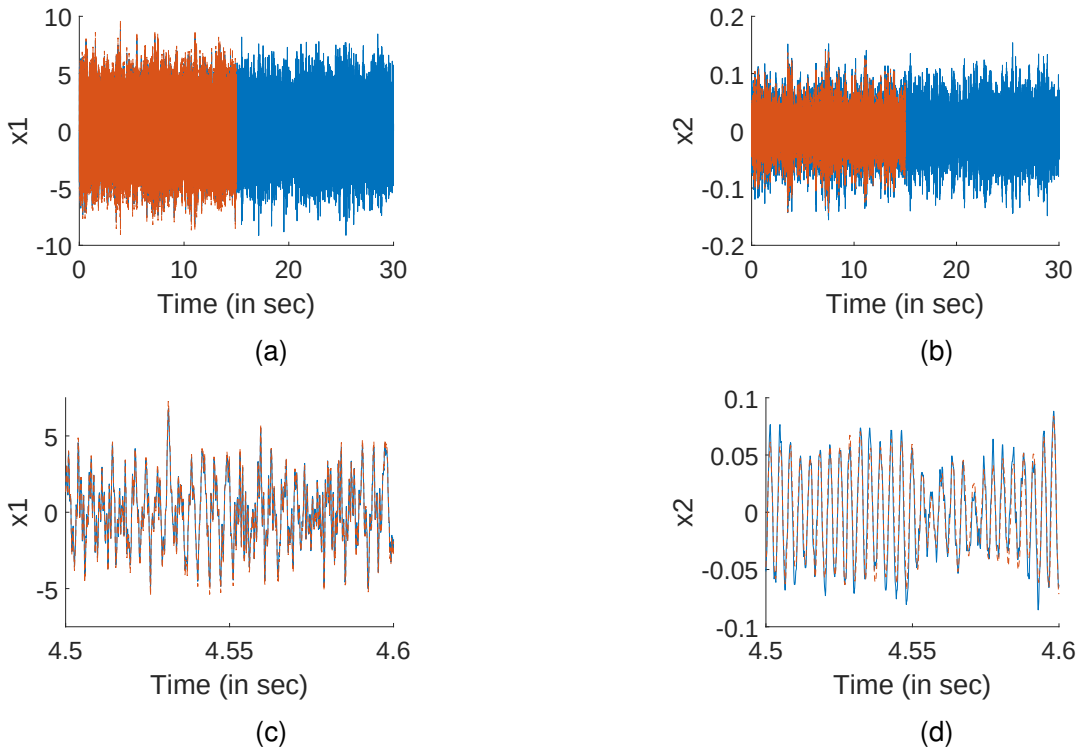


Figure 4.10: Time measurements from the experimental setup (blue line) and identified model (red dashed line) during 25 sec and between seconds 4.5 and 4.6: piezoelectric (4.10a-4.10c) and laser (4.10b-4.10d) sensors

4.2 Active control definition

As seen in Chapter 2 part 1.2.1, many active algorithms are efficient for vibration control. This section focuses on the use of classical feedback control laws as a reference for their ease of use and efficiency.

4.2.1 Applying the control to the structure

A feedback control law $C(s)$ is implemented on the system $G(s)$ which is the smart structure to be controlled, as defined in the block diagram figure 4.11. In this architecture, $r = 0$ is the target signal, e the error, v_d the disturbance, v_c the command, v the input actuation signal of the system $G(s) = [G_1(s); G_2(s)]$ and $[x_1; x_2]$ the respective system outputs. Both the input disturbance and the control signals are applied to the system through one channel v.

The control objective is to minimize the free-beam end vibrations x_2 with two approaches, considering band limited noise v_{d1} ($F_s = 10000Hz$ and $F_c = 1000Hz$) or chirp v_{d2} ($f_0 = 10Hz$ to $f_1 = 5000Hz$ $t_f = 30sec$) disturbance in order to compare with H_2 and H_∞ norm minimization. The feedback control is implemented between the two piezoelectric elements of the beam.
First, a derivative feedback control law is chosen to add damping at low frequencies, such that:

$$C_1(s) = K_d s \frac{N}{s+N} \tag{4.4}$$

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

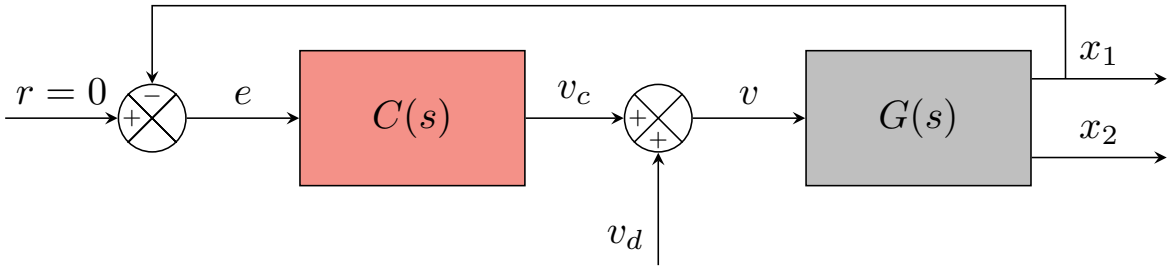


Figure 4.11: Control loop block diagram

Where s is the Laplace variable, $K_d \in \mathbb{R}$ is the derivative gain to tune and $N = 2000\pi$ is the filter coefficient. Then to improve stability and damping on the structure, an other control law $C_2(s)$ is designed with a first order lead compensator to tune, such that:

$$C_2(s) = K_c \frac{s - z_0}{s - p_0} \tag{4.5}$$

Here, $K_c \in \mathbb{R}$ is the gain, $z_0 \in \mathbb{R}$ and $p_0 \in \mathbb{R}$ are the lead coefficients identified as the pole and zero of a transfer function. The objective of this control law is to minimize the vibrations at the free-beam end x_2, by using the lead compensator to increase the stability. To do so, the three parameters of the control law, K_c z_0 and p_0, need to be properly tuned. Note that for both control laws the high frequency limit tends towards the gains K_d and K_c. The roll-off is handled by conditioning and sampling the signals in the model and with the test bench equipment (acquisition system and amplifier).

4.2.2 Parametric tuning

A parametric approach is performed to tune the control laws parameters with criterion minimization based on the closed-loop transfer functions $G_{c,1}(s)$ and $G_{c,2}(s)$, such as:

$$G_{c,1}(s) = \frac{G_1(s)}{1 + G_1(s)C(s)} \qquad G_{c,2}(s) = \frac{G_2(s)}{1 + G_1(s)C(s)} \tag{4.6}$$

The first chosen criterion for the minimization is based on the H_2 norm of the closed-loop transfer function restricted to a finite frequency band:

$$\|G_{c,q}(j\omega)\|_2 = \sqrt{\frac{1}{2\pi} \int_{\omega_1}^{\omega_2} G_{c,q}(j\omega)^2 d\omega} \tag{4.7}$$

with $j \in \mathbb{C}$ the imaginary variable and ω the frequency in $[rad/sec]$ and $q \in [1; 2]$.
The second criterion is defined with the H_∞ norm in the case of linear systems within the same frequency band:

$$\|G_{c,q}(j\omega)\|_\infty = max(G_{c,q}(j\omega)) \tag{4.8}$$

For parametric lead controller tuning, range values are defined as vectors, such as:

$$K_d = \{-0.1 : 0.01 : 0.5\}; A = \{-5000 : 500 : 5000\}; B = \{-5000 : 500 : 5000\}; \tag{4.9}$$

The most efficient controller is found after testing all possible combinations, resulting in 26901 trials for each case. The performance of the tuned controllers are used as a reference to compare with the DRL tuning in parts 4.4.3 and 4.4.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

4.2.3 Disturbance signals

Different disturbance signals have to be considered to obtain equivalent metrics in the time domain. The disturbance signal $v_{d1}(t)$ is a random noise $Rnd(t)$ with a sample frequency F_s in Hz and a second-order filter with a cutoff frequency F_c in Hz, such as:

$$v_{d1}(t) = Rnd(t) * \mathcal{L}^{-1}\{\left(\frac{2\pi F_c}{s + 2\pi F_c}\right)^2\} \tag{4.10}$$

The root mean square (RMS) value of the system outputs, disturbed by this band-limited noise signal, represents the H_2 norm in the time domain.
The signal $v_{d2}(t)$ is a chirp with a sample frequency F_s starting from f_0 to f_1 in Hz at time t_f in seconds, such as:

$$v_{d2}(t) = V_{d2} sin(\phi(t)) \tag{4.11}$$

$$f(t) = \frac{1}{2\pi}\frac{d\phi(t)}{dt} = f_o + \frac{(f_1 - f_0)}{t_f}t \tag{4.12}$$

Where $\phi(t)$ is the phase considering $\phi(0) = 0$ and $V_{d2} \in \mathbb{R}$ is the constant amplitude. The maximum absolute value of the system outputs, disturbed by this chirp signal, is representative of the H_∞ norm in the time domain. In the result section, the metrics are used to compare controlled results.

4.3 Deep reinforcement learning applied to the system

The presented derivative from equation (4.4) and lead from equation (4.5) control laws are now tuned using the DRL algorithm and a TRPO optimization method for neural network (NN) training. To apply the DRL method to a vibration control problem, interactions between the algorithm and environment, driven by Observations, Actions, and Rewards, must be defined. This section delves into the design of DRL components according to this application. The DRL algorithm has to be trained by multiple interactions with the environment to be confident in the parameter tuning of the control law according to the Reward definition. An episode run on the experimental setup is achieved in 1 minute, while it takes only 1 second with the model. Furthermore, tuning parameters that lead to instabilities do not need to be tested on the real structure. Thus, the identified beam defined in Chapter 3, Part 4.1.2.2 is the environment used to train the DRL.

4.3.1 Environment definition

Deep Reinforcement Learning Observation, Action and Reward needs to be defined according to the vibration mitigation objective on the smart structure.

4.3.1.1 Action

The environment is defined by the transfer functions G_1 and G_2, estimated from the experimental measurements (Part 4.1.2.2). The agent only requires a single interaction with the environment before adjusting its policy by updating both the Critic and the Actor neural networks, given the consistent nature of the structural properties that remain unchanged over time and across episodes. The control law applied to the system is

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

also considered part of the environment and its parameters are the actions chosen by the Actor. Constraining Action boundaries to the Actor helps to decrease the NN training time since it has to search for a value within a bounded domain.

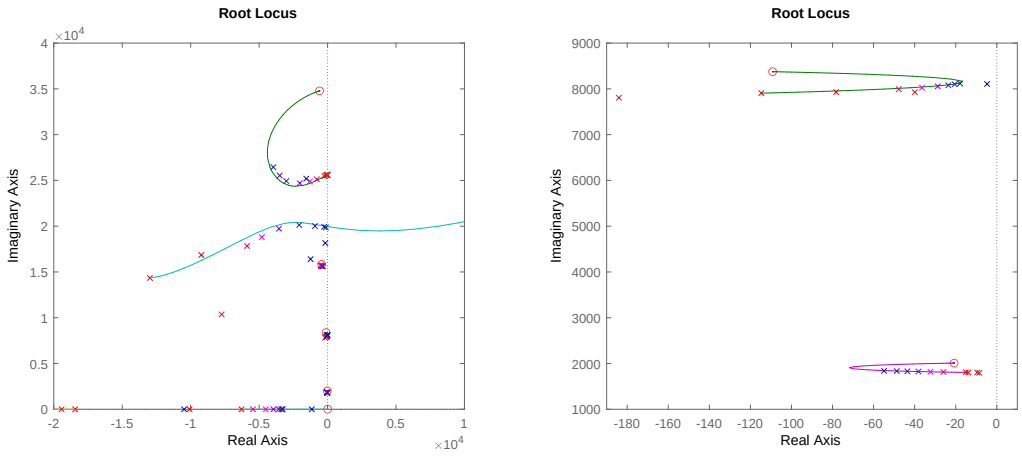
As an action A_t to tune C_1 controller, the RL algorithm is constrained to tune a derivative gain K_d around the stable domain, such as :

- $K_d \in \{-1 \times 10^{-5}; ...; 7 \times 10^{-5}\}$.

To tune C_2 controller, the lead coefficients boundaries are defined according to the sampling frequency F_s (Hz) of the signal and the Shannon criterion, such as:

- $z_0 \in [-F_s\pi; F_s\pi]$
- $p_0 \in [-F_s\pi; F_s\pi]$

Stability is ensured when the real part of the poles of the closed loop is negative (Chapter 2 Part 1.2.1).
As an example, figure 4.12 displays the root locus of the closed-loop transfer function model $G_{c,1}$ between the piezoelectric actuator and sensor with a lead controller, where $z_0 = 0$ and $p_0 = 6280$, based on the identified poles and zeros. This configuration is identical to the derivative controller case. To compare with the experiment, poles and zeros are identified from measurements using the same controller and K_c variations from $K_{c_{min}} = -0.04$ to $K_{c_{max}} = 0.5$. With these z_0 and p_0 lead parameters, stability is



(a) (b)

Figure 4.12: Root locus from identified controlled model $G_1C(s)$ open loop with (a) $z_0 = 0$ and $p_0 = 6280$ and (b) a zoom around the two first poles and zeros (o) poles ($\times$) estimated from measurement on the experimental setup when control is on for different values of K_c from $K_{c_{min}} = -0.04$ ($\times$ dark red) to $K_{c_{max}} = 0.5$ ($\times$ dark blue)

guaranteed for $K_c < 0.6$. However, experimental measurements show that the system becomes unstable when $K_c > 0.5$. The model overestimates the range of K_c values that lead to a stable system. Additionally, it does not consider instability coming from the second pole and creates a fictive zero with a negative real part in the high frequencies above the range of interest. Considering different lead controllers, the system instability occurs for different values of K_c. In order to include all gains leading to a stable system, the lead boundaries are set with:

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

- $K_c \in [-1; ...; 1]$.

This range is larger than the equivalent derivative case and is a challenge for the DRL algorithm to work with unstable controller configurations.

4.3.1.2 Observation

The sensors on the defined system provide time data observations from the environment. In the context of vibration mitigation, it is not necessary to forecast time data because using each time value of the signal will not accurately represent the system's overall behavior. Optimization of the system will not occur at each time step.
For the derivative controller tuning, the average value over time of x_1 and x_2 absolute values are environmental observations S_1 and S_2. Alternatively for the lead tuning, significant metrics can be employed to define environmental observations, such as the root mean square (RMS) of the system's output values during 30 seconds when the disturbance signal is v_{d1}:

$$\begin{bmatrix} S_1 \\ S_2 \end{bmatrix}_{v_{d1}} = \begin{bmatrix} rms(x_1) \\ rms(x_2) \end{bmatrix} \tag{4.13}$$

Considering the chirp disturbance signal during 60 seconds v_{d2} as input, the maximum of the system's output time values are used as Observations:

$$\begin{bmatrix} S_1 \\ S_2 \end{bmatrix}_{v_{d2}} = \begin{bmatrix} max(|x_1|) \\ max(|x_2|) \end{bmatrix} \tag{4.14}$$

These metrics allow to reduce the amount of data input into the agent without losing the meaning of the information on state of the vibrating system. They correspond to H_2 and H_∞ norms in time-domain.

4.3.2 Reward

The objective of vibration mitigation is given to the DRL algorithm through a definition of the Reward function. DRL algorithm adapts its policy in order to increase a numerical Reward which is computed based on the environment behavior after applying the chosen control parameters.

4.3.2.1 Equation

In the objective of minimizing vibrations at the beam-free end, the reward $R_0 \in \mathbb{R}$ is based on the laser measurement x_2, considering the input disturbance v_{d1}:

$$R_0 = \frac{1}{10} \sum_{i_r=1}^{10} \left(\frac{1}{\sum_{i_k=i_1}^{i_2} x_2(i_k)^2} \right) \tag{4.15}$$

Where $i_1 = 0.9(i_r - 1)N_y + 1$, $i_2 = 1.1 \times i_r \times N_y$, $10 \times N_y \in \mathbb{N}$ is the signal x_2 length and $i_r \in \mathbb{N}$ is signal sample index. This formulation creates an overlap with a sliding window on the data measured for 30 seconds.
Figure 4.13 shows reward variations with derivative gain K_d. The reward function is computed with time dependence to account for control performance and stability, avoiding

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

x_2 divergence with time. According to the reward function, the optimal parameter estimation achieves the maximum value of the reward. However this definition does not take into account system instabilities introduced by the derivative feedback loop (Chapter 2 Part 1.2.1).

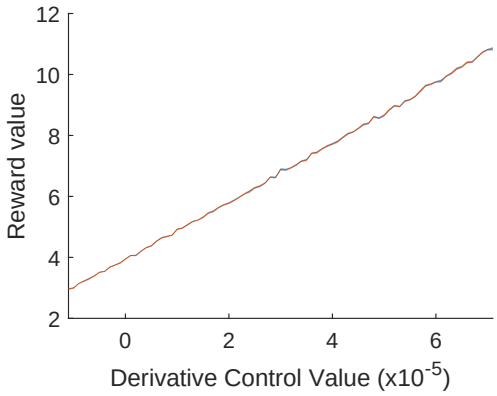


Figure 4.13: Average reward values variations with derivative K_d control values (red line) and the 0.95 confidence interval (blue area) for 10 executions.

For lead controller tuning, rewards are defined to address instability concerns. Considering the input disturbance v_{d1}, reward $R_1 \in \mathbb{R}$ inspired by the H_2 norm definition in the time domain is defined as:

$$R_1 = 20 \times log_{10}\left(\frac{rms(x2_{off}(t))}{rms(x2_{on}(t))} + \frac{rms(x1_{off}(t))}{rms(x1_{on}(t))}\right) \tag{4.16}$$

Where $[x1_{off}; x2_{off}]$ are output time signals when the system is not controlled and $[x1_{on}; x2_{on}]$ are output time signals when the system is controlled.

Based on H_∞ norm definition in the time domain, with the chirp input disturbance v_{d2}, Reward $R_2 \in \mathbb{R}$ is defined as:

$$R_2 = 20 \times log_{10}\left(\frac{max(|x2_{off}(t)|)}{max(|x2_{on}(t)|)} + \frac{max(|x1_{off}(t)|)}{max(|x1_{on}(t)|)}\right) \tag{4.17}$$

To fit the experimental setup, two conditions are added to the two previous rewards:

- The first condition constrains the piezoelectric sensor to an output voltage between -10 and 10 V
- The second condition is a restriction on the gain margin which must be less than 5 dB. Gain margins are computed as exposed in Chapter 2 part 1.2.1.

For all of these cases, Rewards are set to $R_{1,2} = -1$ if the aforementioned conditions are not satisfied.

Figure 4.14 shows the variation of R_1 and R_2 with the gain K_c for different values of (z_0,p_0) tested on the model and a time signal lasting 30 seconds for v_{d1} and 60 seconds for v_{d2}. It reveals the instability of the system when the reward begins to decrease: as the system approaches instability, x_1 progressively increases with K_c until it diverges. The Reward definitions introduce knowledge of control instability to the Reinforcement Learning algorithm by using a penalty value equals to -1. The curves of R_2 are less smooth

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

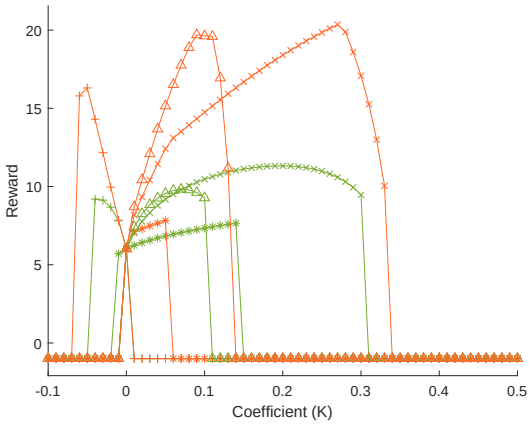


Figure 4.14: Reward R_1 (green line) and R_2 (orange line) variations with K_c for different lead filter z_0 and p_0 combinations : $z_0 = 6280$, $p_0 = -6280$ ($\times$); $z_0 = -31400$, $p_0 = -6280$ ($+$); $z_0 = 31400$, $p_0 = -12560$ ($\triangle$) and $z_0 = -6280$, $p_0 = -21980$ ($*$) tested on the model

than those of R_1. This phenomenon is a result of the reward definition that considers the maximum amplitudes of the time signals and the velocity of the chirps. If the chirp velocity is different, the change in slope will occur at different values of K_c. The energy is concentrated on each frequency contained in the chirp signal, and the movement toward instability could affect other frequencies that are not excited. This effect is observed when the system is truly unstable.

Here, different rewards are defined to observe the AI's ability to adapt to different complexity of functions: linear in the case of R0 and convex non-linear in the case of R1 and R2. The definition of the rewards has been considered according to the disturbance signals in order to best represent the vibrational state of the structure.

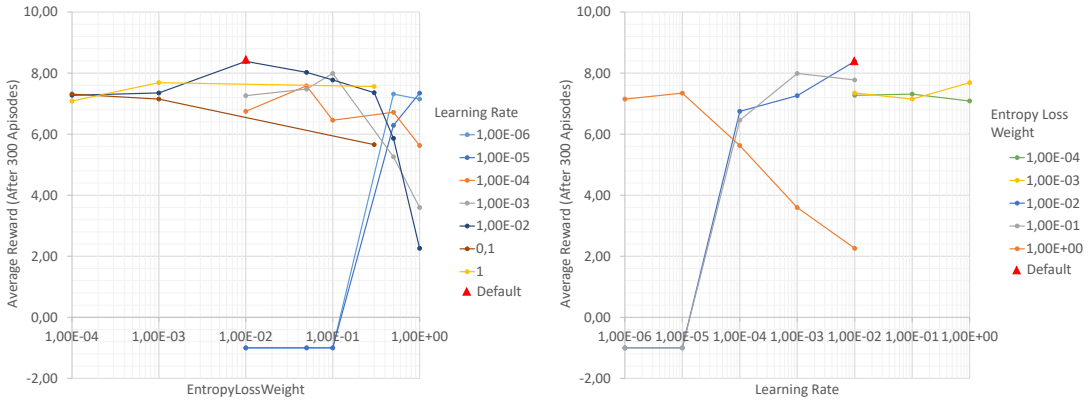
4.3.3 Algorithm

In this section it is chosen to focus only on the use of TRPO algorithm described in Chapter 2 Part 2.2.2. Using this method requires tuning several parameters and adjusting the dimensions of the actor and critic neural networks.

4.3.3.1 Hyperparameters

The hyperparameters of the algorithm govern the training process. For example, increasing the weight of entropy loss will promote exploration and can avoid staying at local minima. The best tuning for these parameters may change according to the environment, but using the default parameters will fit in many cases (Chapter 2 part 4.2.2 Table 10). Figures 4.15a and 4.15b show respectively the effect of entropy loss weight and learning rate on the average reward after 300 episodes of a TRPO agent with a 1-layer, 5 neuron NN size. Each point on the figure corresponds to the training of the same initialized NN. However, a small learning rate can lead to a drop in training performance and significantly increase the training time needed to converge. To tune the controllers, parameters are chosen close to the default ones and detailed in Table 4.5.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés



(a) (b)

Figure 4.15: Training influence with hyperparameter variations: (a) entropy Loss weight and (b) learning rate

Table 4.5: TRPO algorithm characteristics

Name	Symbol	Value
Experience Horizon	N	1
Mini Batch Size	M	1
Entropy Loss Weigth	w	0.01
Number of Epoch	k	1
Average Estimate Method		"gae"
GAE Factor	λ	0.95
Conjugate Gradient Damping	δ_g	0.1
KL-Divergence Limit	δ	0.01
Number Iteration Conjugate Gradient	N_{Cg}	10
Number Iteration Line Search	n	10
Conjugate Gradient Residual Tolerance	C_{gr}	1e-8
Normalized Advantage Method		"none"
Advantage Normalizing Window	N_a	-
Learning Rate	α	0.01
Gradient Threshold	N_{th}	inf
Gradient Threshold Method		"l2norm"
L2RegularizationFactor	l_2	1e-4
Training Algorithm		"adam"
Sample Time	t_s	1 (Event Base)
Dicount Factor	γ	0.99

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

4.3.3.2 Neural network sizing

TRPO method requires the NNs to be adjusted according to the case of vibration mitigation considering cases of continuous actions. Unlike the discrete case, it is not necessary to define all possible actions, only the boundaries between the maximum and minimum values. The output layer size is thus the number of actions to be performed on the environment and the input layer is the number of observations from the environment. This results in smaller NNs, reducing the necessary training time.

Sizing a neural network consists of adjusting the number of hidden layers and the number of neurons per layer. The goal is not to find an optimal NN, but to find a trade-off between computation time and convergence. Overestimating the NNs dimension can result in null weights and biases, redundant neurons, and increased computation time to update all parameters.

Underestimating the dimensions of the neural networks can cause them to struggle to learn properly due to limited memory capacity, increasing the risk of never converging. Researchers are attempting to address this issue by studying how neural networks learn through episodes [55] and how their design affect this process [32] considering the influence of the activation function, the number of layers, the number of neurons per layer, the input metrics and the training parameters.

To illustrate this behavior in the cantilever case, Figure 4.16 shows the effect of NN size on the average reward R_1 after 600 episodes. These results are obtained after training 10 NNs with different initial configurations. The mean and the standard deviation (STD) of the average obtained reward after 600 training episodes show the efficiency and the consistency of the NN size for a given training time and default hyper parameters.

Because TRPO improves monotonically the reward with episodes, training is not terminal for all NNs after 600 episodes, but provides a significant understanding of the expected long-term training behavior. These results highlight that increasing the number of layers improves consistency by reduces STD, but does not necessarily improve performance with the mean reward, especially when the number of neurons per layer is low (example 5 neurons per layer). Obviously, increasing the number of layers increases the number of neurons in the networks, which explains why the NN is efficient with 4 layers and 5 neurons per layer.

The repartitioning of the neurons still has an effect on the training process. For example, using 4 layers with 25 neurons per layer gives more efficient and consistent rewards than using 2 layers with 50 neurons per layer with the same number of neurons. Once the NN is large enough to store all the information needed to model the data, increasing the number of layers and neurons does not change the training performance. For example, considering 50 neurons per layer, the number of layers has little effect on the training results. Some weights and biases may not be used and may be set to zero during the training process. The NN used in this case are relatively small, so the differences in training time variations are only a few seconds.

The Critic and Actor NNs dimensions used for this case are specified in Table 4.6. The dimensions of the NNs have been adjusted according to the previous results.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

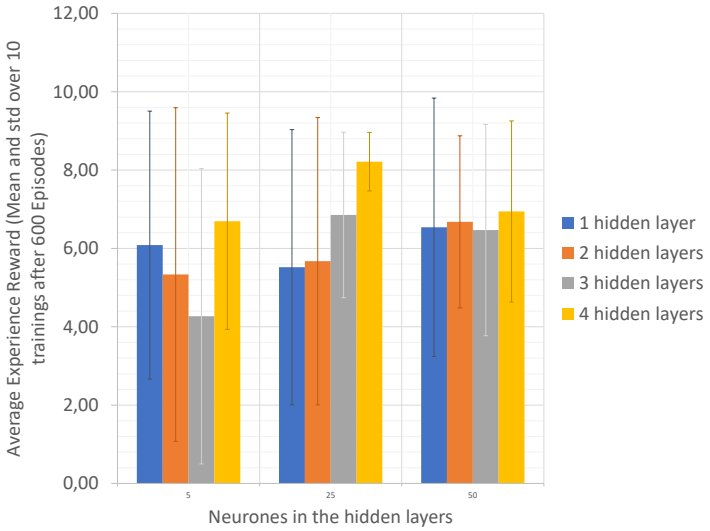


Figure 4.16: Reward R_1 after 600 episodes, mean and standard deviations over 10 different NN initialization.

Table 4.6: Actor and critic neural network definition

Layer	Name	Number of Neurons	
		Actor	Critic
Input	N_i	2	2
Hidden Layer 1	N_{hu}	25	25
Hidden Layer 2	N_{hu}	25	25
Hidden Layer 3	N_{hu}	25	25
Hidden Layer 4	N_{hu}	25	25
Output	No	2	1

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

4.4 Results and performances

To ensure consistency of the method, 10 Agents are trained on the model for the two disturbance cases. Each Agent is initialized with random weights θ and bias ϕ for the input and hidden layers. The TRPO NN optimization method requires starting within a trust region when the structure is stable. One known stable solution is the passive system ($K_c = 0$). To guarantee this constraint, the Actor NN outputs are initialized with output layers weights and bias set to zero. Algorithms run until convergence of all agents is guaranteed to be close to the maximum reward.

4.4.1 Derivative controller

4.4.1.1 Training to tune a derivative controller

The tuning of the derivative control gain is performed taking into account the v_{d1} input disturbance. Figure 4.17 and Table 4.7 show the different training results.

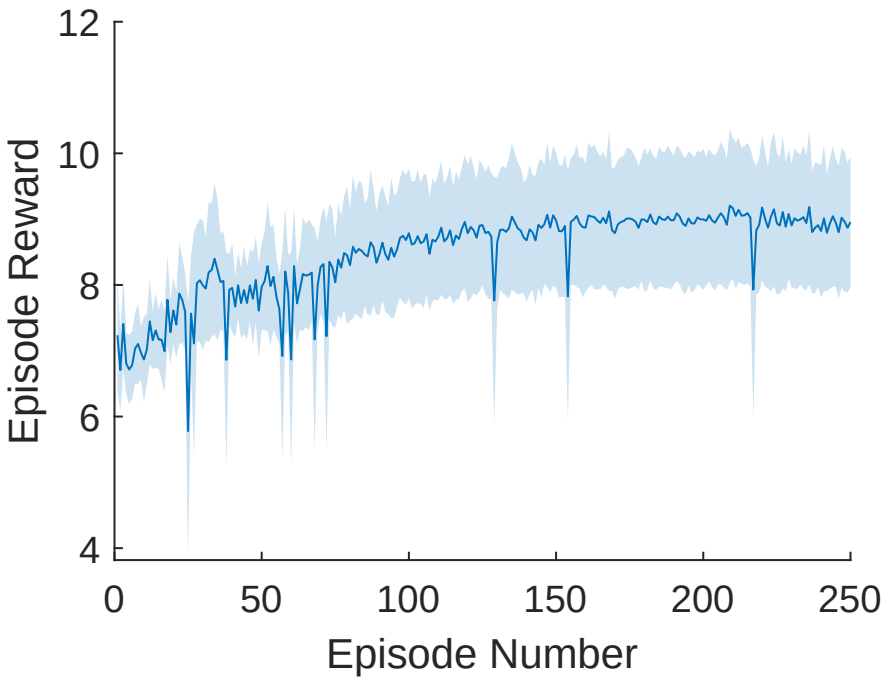


Figure 4.17: Average reward value variations with episode number (dark blue line), the 0.95 confidence interval (blue area) to tune a derivative controller (10 Trainings).

At the end of training, the algorithm maximizes the reward and finds an optimal control value according to the defined action boundary. The K_d value estimated by the different trial runs does not lead to the same value after 250 episodes because of the randomness of the process. Some unfinished training lead to a local minimum. The K_d value with the best performance is the one with the best training end for both the real R_0 and the expected R_{V0} rewards. The number of episodes must be increased to achieve convergence of all agents.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Table 4.7: Training results: derivative controller tuning

Training	Real reward R_0	Expected reward R_{V0}	Absolute error reward	$K_d \times 10^{-5}$	$STD(K_d) \times 10^{-5}$
1	11.40	9.89	1.5046	7.00	0.7389
2	6.80	6.93	0.1262	3.03	0.0041
3	7.60	8.11	0.5085	3.75	0.0007
4	8.92	8.63	0.2922	4.82	0.0005
5	10.63	10.36	0.2746	7.00	0.1055
6	9.19	9.35	0.1603	5.32	0.0018
7	**10.73**	**10.76**	**0.0351**	**7.00**	**0.1231**
8	7.85	7.29	0.5479	4.10	0.0012
9	7.13	7.53	0.4016	3.32	0.0005
10	11.43	10.23	1.1975	7.00	0.5041
Mean	9.37	8.59	0.8441	5.28	0.3121

4.4.1.2 Damping effect on the test bench

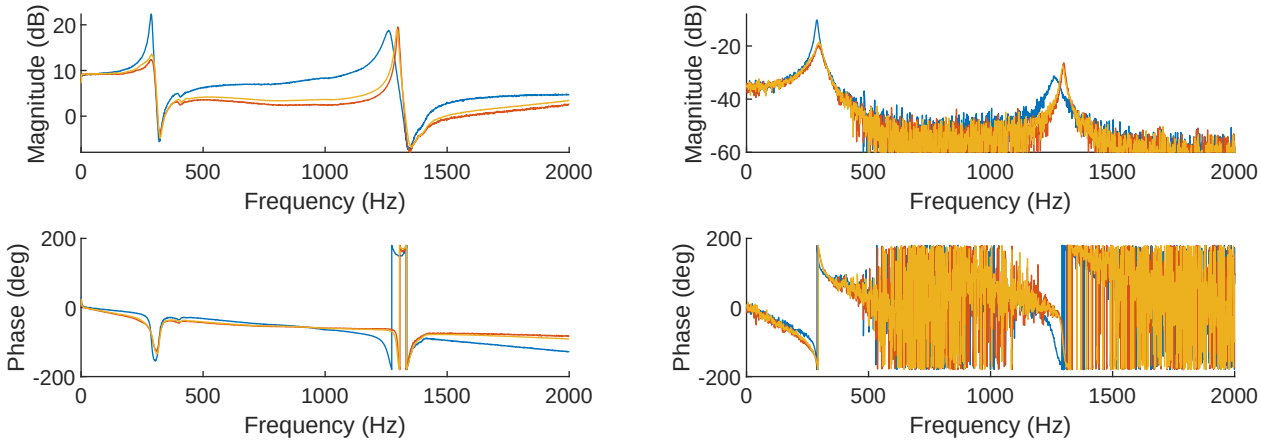
Table 4.8 presents the K_d values obtained with the best training and an average with all operations. Measurements are made on the experimental setup to estimate control performances. Figures 4.18a and 4.18b evaluate transfer functions $G_1(s)$ and $G_2(s)$ with control off and on. Figure 4.18a shows a decreasing phase implying a small delay in the experimental setup. However, the algorithm still finds an efficient and stable control value.

Table 4.8: Derivative control law values and performances.

Method	$K_d \times 10^{-5}$	RMS_{y_2} [mm]	$\|G_2\|_2$	Δ_{dB}
None	0.00	0.0020	10.76	-
$\|H\|_2$	7.00	0.0013	5.93	-10.6
$TRPO_{avg}$	5.29	0.0014	6.59	-9.6
$TRPO_{best}$	7.00	0.0013	5.93	-10.6

$\|H\|_2$ minimization and RL tunning methods lead to an attenuation of 10 dB at 288 Hz at the free-beam end with respect to the first mode as in figure 4.19.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés



(a) (b)

Figure 4.18: Transfer functions (a) $G_1(s)$ and (b) $G_2(s)$ with control off $Kd_{None} = 0$ (blue), control on with $Kd_{TRPO_{best}} = 7 \times 10^{-5}$ (red) and control on with $Kd_{TRPO_{avg}} = 5.29 \times 10^{-5}$ (orange) computed from x_1 and x_2 experimental measurements.

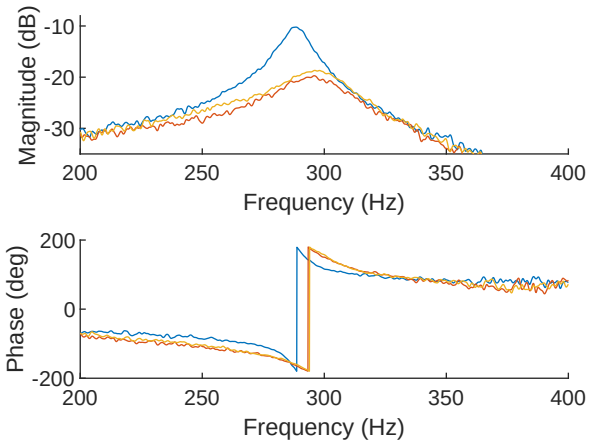


Figure 4.19: Figure 4.18b zoom between 200 and 400 Hz.

These results are promising since the DRL tuning methods succeed in tuning a controller as efficient as the target one obtained with $\|H\|_2$ minimisation. Not all DRLs converged to the same solution. The convergence of different initialised NNs will be studied in the following chapters. The derivative controller only damped the first mode. A spillover effect starts to appear on the second frequency mode (Figure 4.18a). As the algorithms were trained within the stability range of the derivative value, stability problems did not arise. To avoid this, a lead controller is implemented on the structure in the following section. The efficiency of the DRL method is challenged on more complex controllers with multiple parameters to tune and a domain that is no longer unconditionally stable.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

4.4.2 Lead Controller

4.4.2.1 Training dependency to the reward definition to tune a lead controller

Lead controller is tuned considering two cases of disturbances v_{d1} and v_{d2} and their corresponding reward R_1 and R_2 defined in Part 4.3.2.

Figures 4.20 and 4.21 show the average reward variation through episodes. For comparison, two indicative rewards are added to the figures: the reward for the uncontrolled structure and the reward for the structure with the reference controller tuned by parametric H_2 and H_∞ norm minimization according to the disturbance signal. The figures show an overall tendency for rewards to increase with episodes, starting from a reward close to the uncontrolled case and moving above the referenced control case. Due to the reward definition, the NN initialization and a part of randomness in the NN training process, the final maximum Reward reached after at the end of the training can vary. Also, the training progress is not exactly the same for all the Actors.

Considering the disturbance case with v_{d1}, downward variations can be observed during the training (Figure 4.20). These variations can be explained by two main causes. First, little variations of the control parameters can turn the system from stable to unstable due to the reward definition or proximity to the unstable zone with loss of damping. Second, over-fitting in the training can occur if the maximum reward has already been reached and every other tuning can only lead to lower rewards. This over-fitting hypothesis is justified since the average reward stops increasing after 10000 episodes.

Regarding the chirp disturbance v_{d2}, the average reward tendency is linear after 2000 episodes and there are fewer downward variations (Figure 4.21). As shown in Figure 4.14, the training reaches a maximum expected reward of approximately 20. The agent did not appear to have reached a horizontal asymptote after 3×10^4 episodes, but training was stopped because it had reached and exceeded the target reference value.

Table 4.9 and 4.10 show the DRL control law tuning after 20000 and 30000 episodes respectively. Table 4.9 shows that all 10 training sessions lead to a reward above the parametric one. However, two of the sessions (Training number 7 and 10) led to a different controller design and a lower reward compared to the others. This phenomenon highlights a limitation of the DRL method as it may only reach a local maximum. This limitation is also evident in Table 4.10 for training number 1. In this case, the reward range at the end of the DRL training is closer to the reward corresponding to the parametric controller applied to the structure. Based on the reward definition, DRL tuning yields better performance on average than parametric norm minimization tuning for the two tested input signal disturbances on the modeled beam. In the next section, the mitigation performance of the tuned controllers will be examined.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Table 4.9: Training results after 20000 episodes and control law values and performances on model: v_{d1} case.

	R_1	R_{V1}	$\lvert R_1 - R_{V1}\rvert$	$K_{c_{opt}}$	$z_{0_{opt}}$	$p_{0_{opt}}$	RMS(x_1)	RMS(x_2)	$\lVert G_{c,1}\rVert_2$	$\lVert G_{c,2}\rVert_2$
Control off	6.02	-	-	0	0	0	1.96	0.03	4.21	0.03
Parametric										
min($\lVert H\rVert_2$)	10.59	-	-	0.16	6280	-9425	1.89	0.0130	3.01	0.02
Training DRL										
1	11.46	11.31	0.15	0.21	5255	-6231	1.52	0.0125	2.63	0.0138
2	11.49	10.96	0.53	0.23	5332	-6257	1.55	0.0123	2.69	0.0135
3	11.41	11.14	0.28	0.27	5004	-5902	1.66	0.0121	2.85	0.0134
4	11.46	10.89	0.57	0.26	5266	-6232	1.61	0.0122	2.78	0.0134
5	11.50	11.34	0.15	0.25	5202	-6118	1.58	0.0122	2.76	0.0134
6	11.50	11.36	0.14	0.23	5269	-6237	1.55	0.0123	2.68	0.0136
7	11.20	11.21	0.01	0.07	7488	-2580	1.62	0.0127	2.91	0.0140
8	11.50	11.49	0.01	0.24	4960	-5852	1.59	0.0121	2.76	0.0135
9	11.51	11.02	0.49	0.23	5106	-6019	1.56	0.0122	2.70	0.0135
10	11.23	11.22	0.01	0.10	6345	-3407	1.63	0.0125	2.83	0.0139
Mean	**11.43**	**11.19**	**0.23**	**0.21**	**5523**	**-5484**	**1.59**	**0.0123**	**2.76**	**0.0136**
STD	**0.11**	**0.18**	**0.21**	**0.06**	**751**	**1266**	**0.04**	**0.0002**	**0.08**	**0.0002**

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Table 4.10: Training results after 30000 episodes and control law values and performances on model: v_{d2} case.

	R_2	R_{V2}	$\lvert R_1 - R_{V1}\rvert$	$K_{c_{opt}}$	$z_{0_{opt}}$	$p_{0_{opt}}$	MAX(x_1)	MAX(x_2)	$\lVert G_{c,1}\rVert_\infty$	$\lVert G_{c,2}\rVert_\infty$
Control off	6.02	-	-	0	0	0	6.7258	0.1193	43.10	0.999
Parametric										
min($\lVert H\rVert_\infty$)	20.35	-	-	0.270	6280	-6280	2.5796	0.0153	16.46	0.0757
Training DRL										
1	20.34	20.23	0.11	0.18	11602	-9422	1.54	0.020	7.72	0.096
2	20.50	20.41	0.08	0.19	10263	-9663	1.39	0.021	6.97	0.101
3	20.51	19.58	0.92	0.20	9997	-9579	1.41	0.020	7.08	0.101
4	20.49	18.94	1.55	0.20	10159	-9608	1.42	0.020	7.12	0.101
5	20.50	19.35	1.14	0.20	10213	-9684	1.40	0.021	6.99	0.103
6	20.50	20.46	0.04	0.20	10049	-9498	1.42	0.020	7.13	0.101
7	20.50	19.38	1.12	0.20	10003	-9592	1.44	0.020	7.19	0.100
8	20.49	20.51	0.02	0.20	10291	-9682	1.40	0.021	7.02	0.103
9	20.40	19.28	1.12	0.20	10809	-9802	1.58	0.019	7.89	0.094
10	20.39	19.27	1.12	0.20	10709	-9702	1.57	0.020	7.67	0.092
Mean	**20.46**	**19.74**	**0.72**	**0.20**	**10420**	**-9633**	**1.46**	**0.020**	**7.30**	**0.099**
STD	**0.06**	**0.56**	**0.56**	**0.007**	**484**	**115**	**0.07**	**0.0005**	**0.36**	**0.003**

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

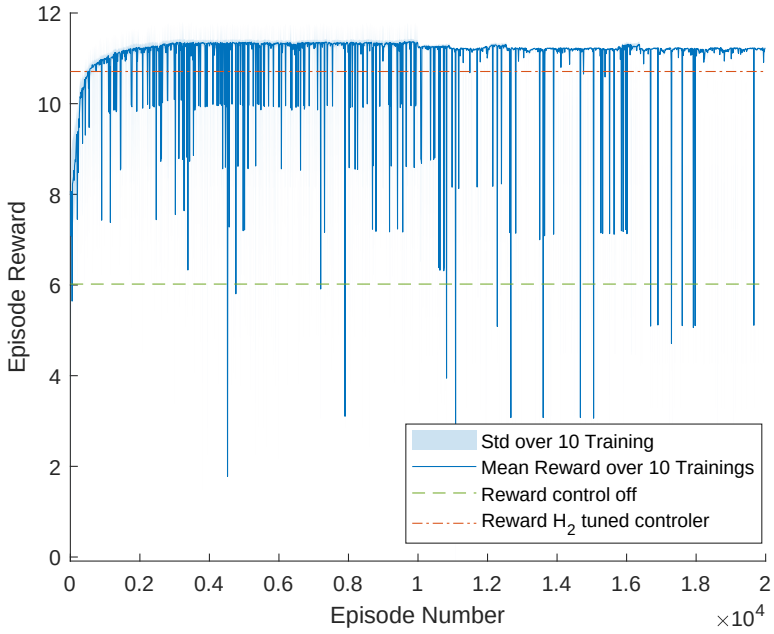


Figure 4.20: Training on model with disturbance v_{d1}: mean reward over all training (dark blue), the corresponding standard deviation area at each episode step (light blue), the equivalent reward when the control is off (green - -) and the equivalent reward with H_2 norm minimization controller parameters (red .-).

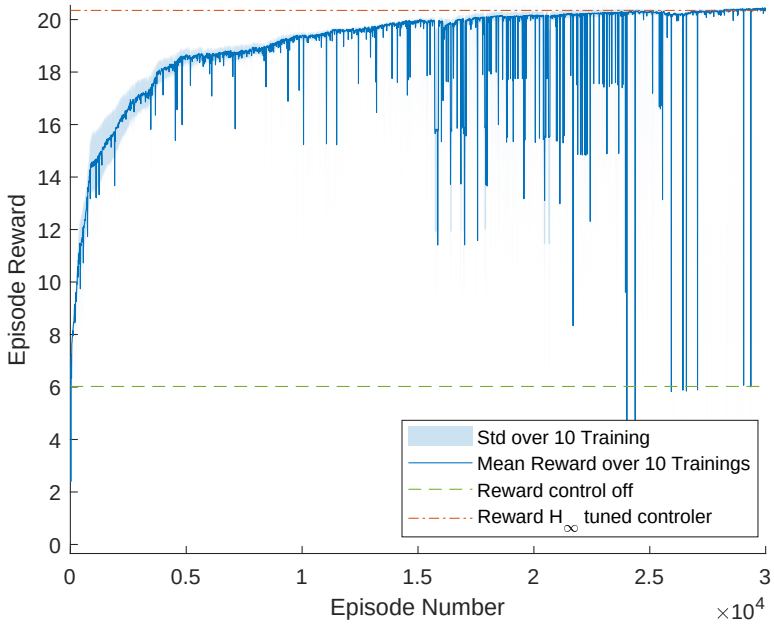


Figure 4.21: Training on model with chirp disturbance v_{d2}: mean reward over all training (dark blue), the corresponding standard deviation area at each episode step (light blue), the equivalent reward when the control is off (green - -) and the equivalent reward with H_∞ norm minimisation controller parameters (red .-).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

4.4.3 Damping effect on the test bench

The model and experimental structure were subjected to previously tuned controllers to assess their effectiveness in reducing vibration, using significant metrics. An indicative reward was computed from the measured data. The performances of each DRL-tuned controller were compared with the control turned off and the parametric-tuned controller in Tables 4.11 and 4.12. Performance of the controller on the model is presented in Tables 4.9 and 4.10.

The DRL and parametric methods give different controller tunings, both of which significantly reduce vibration. Mean values are given to highlight the consistency in terms of control performance given by the DRL tuning method, proving its reliability and consistency.

Figures 4.22 and 4.23 show the effect of the controllers on the frequency response of the modeled and experimental system. These figures show differences between the modeled behavior of the controller and the experimental response of the system, especially at high frequencies.

Considering the v_{d1} disturbance for all DRL tuned controllers, the vibration damping performances in the time domain are very close as observed for $rms(x_1)$ and $rms(x_2)$. They give better damping than parametric tuning for both model and experiment. Within the frequency domain, the differences are more visible. All natural frequencies of the system are damped with all tuned controllers. The difference in controller tuning may be due to the parametric definition of the set of sampled values used for norm minimization. Therefore, multiple controllers lead to similar vibration reduction performances and DRL tuned controllers are more effective than the parametric controller for both model and experimental.

For a chirp v_{d2} disturbance, the DRL tuned controller gives better attenuation than the parametric tuned controller on the model but not on the experiment. In fact, $max(x_1)$ and $max(x_2)$ are lower on the model for the DRL tuned controller compared to the parametric tuned controller, and the contrary occurs on the experimental measurements. These differences are explained by the model underestimating the effect of high frequencies and the static gain of the experimental structure. Also, the reward at the end of DRL training is close to the parametric tuned controller. This explains why the damping performances of the two tuned controllers are close.

Nevertheless, the DRL-tuned controllers show a significant reduction in structural vibrations. The tuned controllers for Case 1 and Case 2 are not the same. This proves the efficiency of the method to tune a controller adapted to different input disturbances.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Table 4.11: Case 1: lead control Law values and performances measured on the experimental setup

Tuning Method	Tuned Values			Time Metric		Norm		Experimental Reward
Case 1: Noise v_{d1}	K_c	z_0	p_0	RMS(x_1)	RMS(x_2)	$\|G_{c,1}\|_2$	$\|G_{c,2}\|_2$	$R1$
None	0	-	-	1.9400	0.039	874.31	8.76	6.02
$\|H\|_2$	0.160	6280	-9425	1.6436	0.023	568.48	4.51	9.13
DRL 1	0.21	5255	-6231	1.38	0.017	540	3.59	11.33
DRL 2	0.23	5332	-6257	1.42	0.017	554	3.49	11.22
DRL 3	0.27	5004	-5902	1.48	0.017	579	3.32	11.25
DRL 4	0.26	5266	-6232	1.48	0.017	584	3.35	11.15
DRL 5	0.25	5202	-6118	1.45	0.017	569	3.36	11.14
DRL 6	0.23	5269	-6237	1.42	0.021	553	3.48	10.19
DRL 7	0.07	7488	-2580	1.46	0.018	587	3.41	10.98
DRL 8	0.24	4960	-5852	1.41	0.018	552	3.37	11.11
DRL 9	0.23	5106	-6019	1.41	0.019	549	3.42	10.86
DRL 10	0.10	6345	-3407	1.42	0.019	556	3.46	10.87
DRL Mean	**0.21**	**5523**	**-5484**	**1.43**	**0.018**	**562**	**3.43**	**11.01**
DRL STD	**0.06**	**751**	**1266**	**0.03**	**0.001**	**16**	**0.08**	**0.33**

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

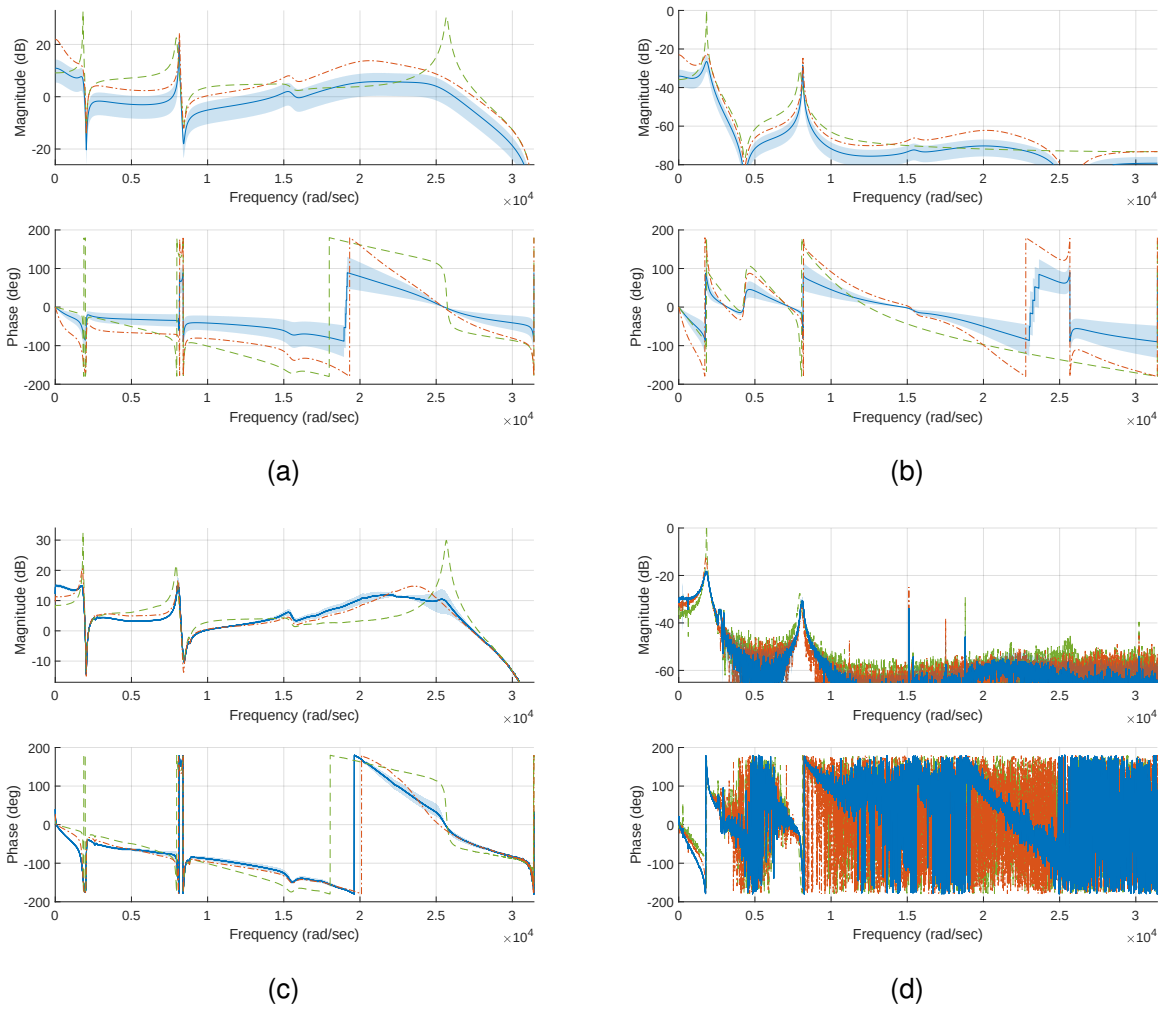


Figure 4.22: Case v_{d1}: system passive response (green - -), system controlled response with parametric H_2 norm minimisation (red .-), mean (dark blue) and standard deviation (light blue) over all system controlled response with DRL tuned controller: (a) model and (c) experiment $G_{c,1}$ bode diagram and (b) model and (d) experiment $G_{c,2}$ bode diagram.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Table 4.12: Case 2: lead control law values and performances measured on the experimental setup

Tuning Method	Tuned Values			Time Metric		Norm		Experimental Reward
Case 2: Chirp v_{d2}	K_c	z_0	p_0	MAX(x_1)	MAX(x_2)	$\|G_{c,1}\|_\infty$	$\|G_{c,2}\|_\infty$	R_2
None	0	-	-	7.25	0.067	42.42	1.13	6.02
$\|H\|_{inf}$	0.27	6280	-6280	2.23	0.043	9.10	0.10	13.62
DRL 1	0.18	11602	-9422	3.30	0.043	11.46	0.14	11.49
DRL 2	0.19	10263	-9663	3.74	0.057	13.11	0.14	9.88
DRL 3	0.20	9997	-9579	4.41	0.040	15.62	0.14	10.43
DRL 4	0.20	10159	-9608	4.67	0.047	16.71	0.13	9.50
DRL 5	0.20	10213	-9684	5.01	0.058	18.06	0.13	8.27
DRL 6	0.20	10049	-9498	4.25	0.069	14.98	0.14	8.57
DRL 7	0.20	10003	-9592	4.47	0.065	15.86	0.13	8.46
DRL 8	0.20	10291	-9682	5.12	0.062	18.62	0.13	7.95
DRL 9	0.20	10809	-9802	5.94	0.053	24.25	0.13	7.94
DRL 10	0.20	10709	-9702	5.54	0.057	21.07	0.13	7.93
DRL Mean	**0.20**	**10420**	**-9633**	**4.65**	**0.055**	**16.97**	**0.13**	**9.04**
DRL STD	**0.01**	**484**	**115**	**0.80**	**0.01**	**3.74**	**0.01**	**1.23**

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

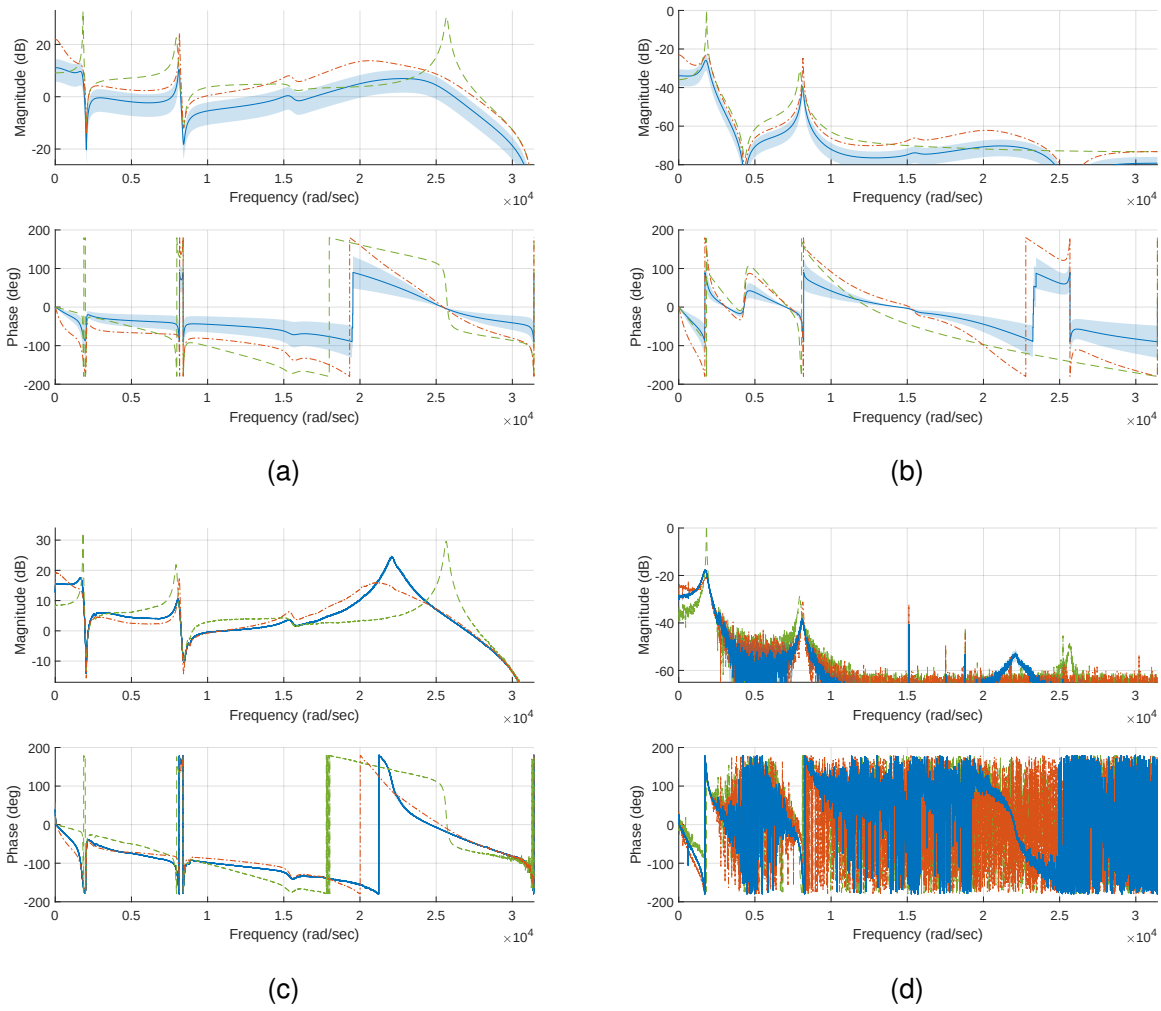


Figure 4.23: Case v_{d2}: system passive response (green - -), system controlled response with parametric H_∞ norm minimization (red .-), mean (dark blue) and standard deviation (light blue) over all system controlled response with DRL tuned controller: (a) model and (c) experiment $G_{c,1}$ bode diagram and (b) model and (d) experiment $G_{c,2}$ bode diagram.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

4.5 Discussion

Using a Deep Reinforcement Learning algorithm to tune a control law requires expertise in both machine learning and vibration control fields. This section is a discussion of the implementation, use and performance of the vibration mitigation method.

4.5.1 Benefits and limits of DRL approach

Reinforcement learning algorithms require trial-and-error computation to train neural networks, which is time-consuming and resource-intensive. In contrast, the number of iterations to perform for parametric tuning depends only on the number of parameters and the degree of accuracy expected (defined by the range of parameter values). To tune a derivative controller, 61 trials are required when using parametric whereas more than 250 to train a NN. With only one parameter to tune, DRL is too sophisticated, but it gets interesting when you consider multiple parameters. For three-component lead tuning, 26901 parameter configurations are tested to determine an efficient reference controller (Part 4.2.2). Considering case one with disturbance v_{d1}, the DRL algorithm succeeds in tuning a controller as efficient as the reference one in less than 10000 episodes. It is less than twice the number of trials required with parametric. Considering the case v_{d2}, the difference between DRL and parametric is less significant since it requires about 27000 training episodes to achieve parametric tuning efficiency. With the same number of trials, using DRL still requires more computational resources to update weights and biases.

For experimental vibration control applications, multiple tests are time-consuming considering the acquisition time and the NNs updating process. Moreover, poorly designed and unmastered neural networks can result in closed-loop instability and electrical issues, requiring human intervention to shut down the system. To ensure the efficiency of a DRL-tuned controller, it is more convenient to use an experimental model for running multiple training on a powerful hardware computation unit. The model must be as close as possible to the experiment in order to account for instabilities and to tune an effective, efficient controller. If the model does not include experimental instabilities, a numerically stable and efficient controller may be tuned but be experimentally unstable when applied on the testbench.

Observations, actions, and rewards define the links between the environment and agent. Meaningful metrics should be used to define observations and actions according to the vibration mitigation objective. It is recommended to normalize the data inputs and outputs when the minimum and maximum values are known to prevent neural network divergence. In this experiment on controlling vibrations, we can estimate action limitations based on system stability. This reduces the domain of parameter exploration and increases training speed.

DRL tunes controller parameters in an autonomous way, but the methods still have their own parameters to tune: Hyperparameters and the dimensions of the agent neural networks. They must be set according to the problem to be solved, but it is not necessary to use optimal ones. As seen in parts 4.3.3.1 and 4.3.3.2, empirical methods are mostly used to tune DRL parameters. Further research should be done to propose a reliable method for tuning training parameters, especially for reinforcement learning algorithms.

Depending on the initialization of the neural network, divergence during the training process may occur. Applying the same algorithm can result in diverse outcomes after an equivalent number of training iterations. This is due to the probabilistic approach of

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

the algorithm that occurs during NN initialization and training. It highlights the impact of randomness on both computation time and achieved results. The randomness of the training process is determined by the numerical seed. It is recommended to train multiple neural networks in case of random initialization and to continue training until the highest reward is achieved. When using TRPO, it is mandatory to initialize the neural network with a stable solution for the closed-loop since the algorithm searches for a solution by expanding the search space around a starting point, which is known as stable. The impact of the random initialization is studied in Chapter 5.

At the end of the training, the DRL method will find a solution, but there is no guarantee that it will reach one optimal solution. It can converge to a local maximum because of overfitting or because the agent needs more episodes to learn. Although this is not the case in this experiment, the end of the training process can be defined by a minimum reward value to reach according to a control performance metric. In this context, the objective is not to find the optimal tuning for a controller, but rather to allow for a trade-off between computation time and control performance. Utilizing DRL to tune a controller enables the creation of an efficient and automatically tuned controller.

4.5.2 Stability concern

The DRL relies on reward definitions to fine-tune the controller. The tuned controllers are not identical because of the reward and the input disturbance signal definition. H_∞ norm is more robust by considering all frequencies minimization. For both disturbances applied to the structure, the tuned control laws present zeros with positive real parts but do not lead to closed-loop instabilities. The use of zeros in the closed-loop transfer function ensures a phase close to 90 at the system's first natural frequency. The tuned controllers tend to soften the system at low frequencies by decreasing the static stiffness. This critical point must be considered to avoid system instability. The controllers are adapted to the source because the reward function definition is adapted to the source.

Finally, the stability of the system depends mainly on the controller definition, but also on the reward definition. If the action domain contains unstable controllers, the DRL algorithm is informed of instability by the reward function. In this case, instability can be observed in the time domain with signal amplitude increasing. Gain margin and phase margin could be used as complementary indicators to be integrated into the reward function. This strategy will be explored in the following chapters.

4.5.3 Algorithm robustness

This first experimental application case does not explore all the common uses of DRL algorithms. The method interest comes from the NNs ability to store a pattern between input observations from the environment and output actions. Because of these non-linear constitutive equations, NNs are able to adapt to a changing environment and event to store more than 1 pattern at a time to tune different controllers. Examples for extending the method are developed in the following sections.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

4.5.3.1 Robustness to environment variations

In this section, the observations are slightly modified to challenge the algorithm's ability to adapt to a modified environment. A given initialised NN is first trained on the case v_{d2} to tune a lead controller (DRL 1). It is then trained further on the case v_{d1}, keeping the memory of the pre-training (DRL 2). In an other way, the given initialised NN is trained directly on the case v_{d1} (DRL 3). Figure 4.24 shows a flowchart of the implemented strategy.

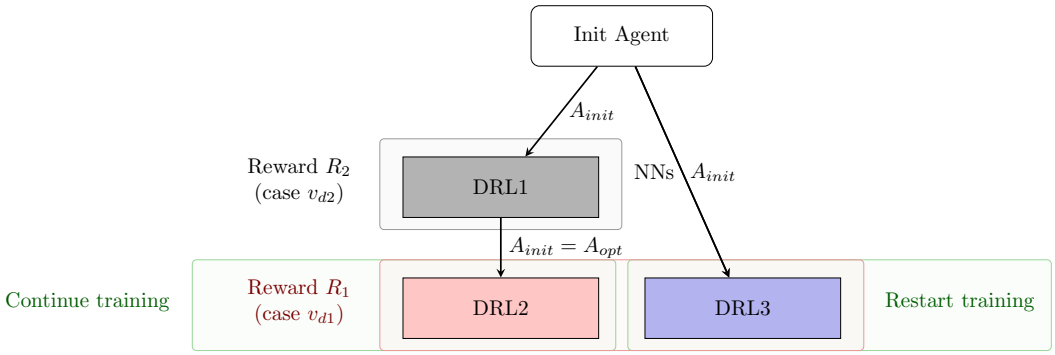


Figure 4.24: Flowchart: restarting or continuing the training to adapt from input noise v_{d1} (Reward 1) to v_{d2} (Reward 2)

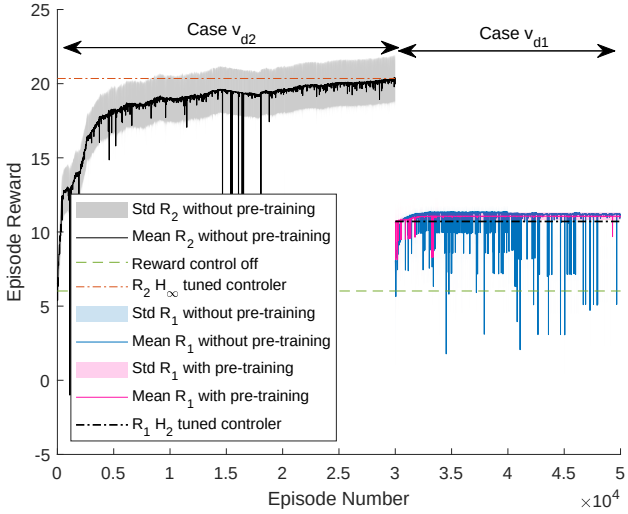
The reward variation with episodes is shown in Figure 4.25:

- Figure 4.25a: A decrease in episode reward value is observed at episode $3 \times 10^4 + 1$ as the environment changes according to the input perturbation, observations, and reward definition.
- Figure 4.25b highlights the differences between pre-trained and non-pre-trained NN. For both training cases, the tuning parameters are close (tables 4.11 and 4.12).

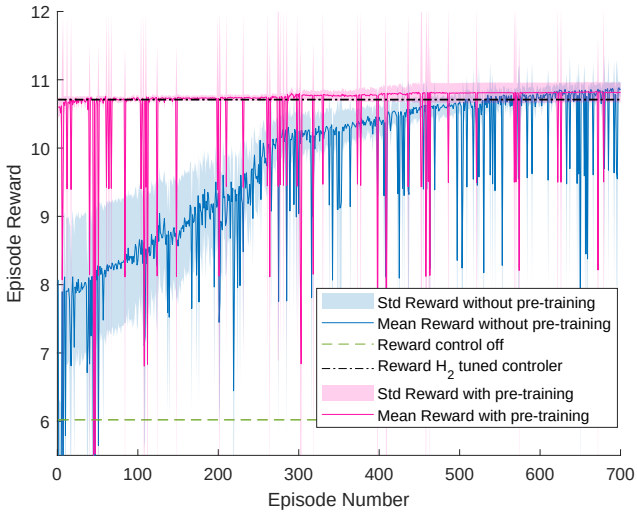
Since TRPO slowly improves parameters, pre-training NN suggests a tuning close to the target tuning even on a different case. Pre-training changes NN's initialization for training, explaining why it goes beyond the reference reward tuned with H_2 norm minimization in less than 10 episodes. After 600 episodes, NNs without pre-training reach the rewards obtained with pre-training and lead to the same controller tuning. These results demonstrate the interest of the proposed method through its ability to adapt quickly to environmental variations thanks to neural networks memory.

To go further, it may be relevant to observe the robustness of the algorithm with respect to mass or stiffness structure modifications. It is also critical to quantify the allowable differences between the trained environment and the modified environment that lead to an improvement in training. This information defines the level of accuracy required for model definition. Differences may occur between the model and the experiment, leading to a less efficient control of the structure vibrations. To address this issue, model can be used to pre-trained NN before training it on the experiment in fewer episodes close to a stable domain.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés



(a)



(b)

Figure 4.25: Training on model with noise disturbance v_{d1} (a) with and without pre-training with chirp disturbance v_{d2} and (b) a zoom on the 700 first episodes of the training on case v_{d1}.

4.5.3.2 Robustness to multi environment conditions

NN output adapts to the observations and is able to learn more than one pattern in parallels. In this section, the two cases, considering v_{d1} v_{d2} disturbances, are learned simultaneously by the same NN to tune the lead controllers parameters. Thus, the NN tunes six parameters simultaneously. To do this, an episode is divided into two epochs as shown in Figure 4.26. The first epoch observes and tunes the controller in case one (v_{d1} and Reward 1) and the second epoch focuses on case two (v_{d2} and Reward 2) .

To update the NN weight and bias, the algorithm sums the reward obtained in each epoch R_1 and R_2. In this way, it improves the global performance considering both cases. With this method, the risk is to improve only one of the two cases, but this phenomenon is avoided after a large number of epochs. Figure 4.27 shows the training of the Agent

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

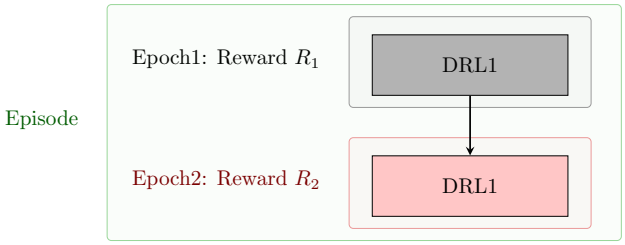


Figure 4.26: Flowchart: multi epoch training to adapt from input noise v_{d1} (Reward 1) and v_{d2} (Reward 2)
at the same time.

while tuning the controllers for two cases.

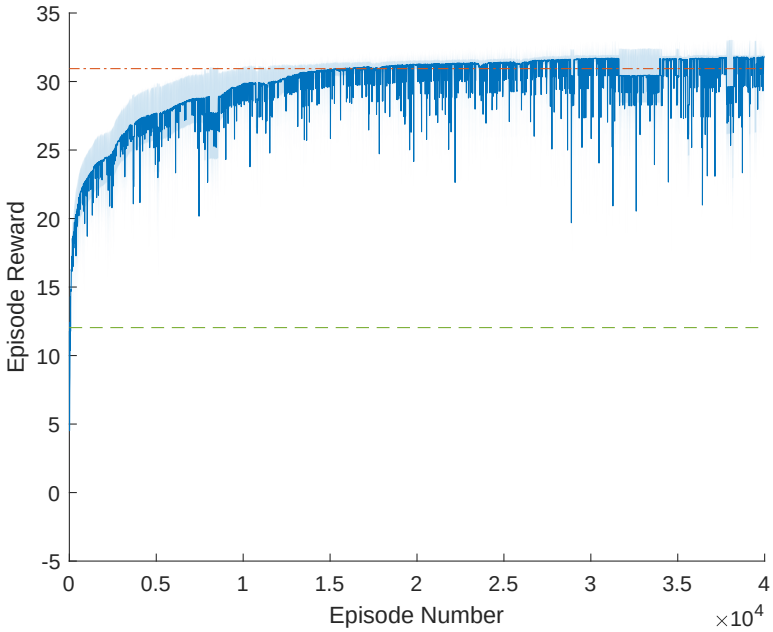


Figure 4.27: Training on model with noise v_{d12} and chirp disturbance v_{d2}: mean reward over all training (dark blue), the corresponding standard deviation area at each episode step (light blue), the equivalent reward when the control is off (green - -) and the equivalent sum of rewards $R_1 + R_2$ with H_2 and H_∞ norm minimization controller parameters (red .-).

Learning on multiple environment definition is possible with NN because of the non-linear math definition of neurons coming from the activation function. Depending on the observations, some neurons are activated others not. Using this property require large enough NN. In this example the network is large enough to be able to learn on two different environments. To go further, it might be relevant to understand how to size the NN to learn a given number of environmental variations simultaneously. In this example, only two environment configurations are learned, but it is well known that NN can learn millions of environment variations with millions of training steps, such as in board games [128], video games [96], or robotics [85]. Training these models requires efficient and costly computational units, which are beyond the scope and resources of this thesis. However, this feature provides the ability to tune controllers to different environmental configurations at the same time. It is not explored in the following chapters, but provides

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

a perspective on the fields, especially when considering non-linear systems.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

4.6 Conclusion

This study shows the DRL algorithm performances in an experimental context for tuning controller parameters to mitigate vibration of a smart cantilever beam.

As a preliminary approach, the TRPO method tuned an efficient derivative gain K_d within the stability domain to minimize cantilever free-beam end vibrations.

As a second step, stability concerns are introduced through the reward definition to tune a three-parameter control law. Two input disturbance signals, namely noise v_{d1} and chirp v_{d2} are used to train the DRL algorithm, leading to different controllers. The DRL tuned controllers are adapted to the input disturbance signals and ensure efficient vibration reduction. When compared to a parametric controller tuning approach that uses norm minimization, the DRL method's efficiency is proven in both disturbance cases.

The training process for DRL can be time-consuming and computationally expensive, depending on the system's complexity. DRL guarantees to find a solution, but it may only reach a local maximum. The definitions of the DRL elements (environment, observation, reward and action) must be properly defined according to the mitigation objective. In this application, the environment consists of the smart cantilever beam controlled by a law whose parameters define the DRL actions. Observations are defined by the RMS or maximum temporal response of the structure in coherence with the input disturbance definition. The reward definition is designed to balance between control performance and structure stability.

A finite element model is used to design the smart structure but building uncertainties introduces significant differences with the experimental setup. Instead, the experimental setup is modeled with measured transfer function estimation. It enables to conduct multiple tests in a short period of time and to fine-tune algorithm and neural network hyper-parameters. By using the model, the DRL algorithm is able to produce the results presented in this chapter in just a few hours, whereas running the algorithm on the actual experiment would have taken weeks. It is important to note that the model used must closely match the experimental system to ensure efficient tuning of the controller.

Further research investigates the robustness of DRL to environmental changes through two examples. They demonstrate the DRL method's ability to adapt to a new environment configuration and to store multiple environment configurations simultaneously.

This method offers possibilities to adapt to complex distributed systems with multi-parameters to tune. In the next Chapter, DRL is tested in more complex systems that include multiple piezoelectric transducers in a network, and distributive control laws to tune.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Chapter 5

Tuning methods for distributed active vibration control of a smart structure : an experimental study

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Deep reinforcement learning algorithms are used to perform distributive control on transportation such as intersection traffic [42], automated vehicles [126], or trajectory design for unmanned aerial vehicles (UAV) [52]. In these examples, multiple vehicles interact with each other in an environment. A parallel can be made with the vibration control application where the vehicles can be seen as the transducers and the environment as the structure to be controlled. In this chapter, deep reinforcement learning (DRL) algorithms are applied to a multi-input multi-output system (MIMO). The study, performed on a smart beam with free boundary conditions, is divided into three parts.

First, the structure with three pairs of colocalized piezoelectric patches is designed using a finite element model to perform modal analysis. The goal is to determine the transducer locations that maximize the piezoelectric coupling coefficient to the target modes below 1000 Hz. An experimental test bench is built and measurements are performed to identify the transfer function system between the actuators and the sensors. A multi-input multi-output model is constructed by pole and zero estimation. The use of a numerical model is essential to perform multiple tests without causing damage to the test bench. The use of test bench measurements rather than finite element modeling results avoids a fitting phase. A positive position feedback (PPF) controller is implemented between each pair to damp the mode vibrations in a specific frequency range between 350 and 400 Hz.

Secondly, simplex and deep reinforcement learning methods are used to tune the parameters of the three controllers. A numerical environment is defined to perform the optimization according to an objective function called the reward. A statistical study is conducted to assess the effectiveness of the methodologies, with the control performances evaluated in a numerical environment.

The methods are tested for robustness by changing the environment, in this case the reward definition. The ability to adapt to variations in the objective function is investigated by extending the target frequency band and improving the stability criteria. New tuned controllers are implemented on the testbed. The results demonstrate the efficacy of deep reinforcement learning tuning techniques in a MIMO context, as evidenced by enhancements in control efficiency, stability, and adaptability. It offers the prospect of automatic tuning, taking into account multiple conditions.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

5.1 Definition of a smart multi-input multi-output (MIMO) system

A structure is designed with distributive transducers working as a network. As a multi-modal structure, the goal is to have the transducers in a position where they have authority over the that need to be controlled. A finite-element model is used to design the structure and perform the modal analysis with COMSOL Multiphysics simulation software. A test bench is then built and a pole-zero transfer function model is constructed from the measurements. The differences between the model and the test bench are highlighted by the effect of the PPF controller on the damping of the target mode.

5.1.1 Design of a smart beam

5.1.1.1 Finite element model definition

The smart structure studied in Figure 6.4 consists of a steel beam of length L_b, width b_b and thickness h_b. Three identical pairs of piezoelectric transducers (PZT-PIC255) of length L_p, width b_p and thickness h_p are bonded to the beam. The top transducers Pa_1, Pa_2 and Pa_3 are used for actuation and Ps_1, Ps_2 and Ps_3 for sensing. All transducers are centered on the beam width. One pair of transducers is centered on the length with the position e_{p2} from the side of the beam. The position of the two lateral pairs of transducers is defined by the distances to the beam ends e_{p1} and e_{p3}. The model dimensions and material properties are defined in Table 6.3. The beam ends are free and the electrical ground is selected on the electrode facing the beam. The finite element model uses a default mesh of 8609 elements.

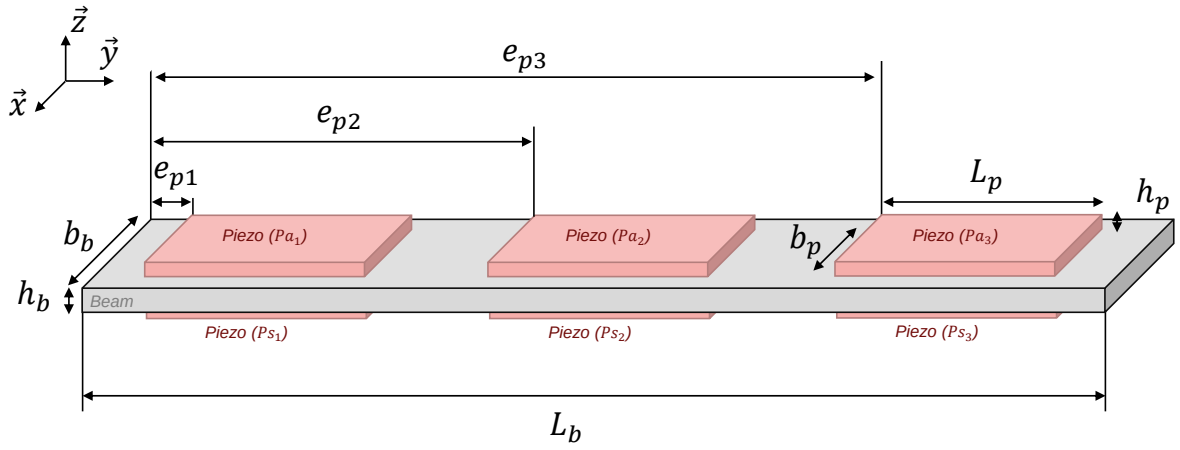


Figure 5.1: Beam dimensions diagram

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Table 5.1: Definition of the constants

	Variable	Beam		Piezoelectric (PZT-PIC255)	Transducer	Unit
		Name	Value	Name	Value	
Dimensions	Length	L_b	400	L_p	61	$[mm]$
	Width	b_b	60	b_p	35	$[mm]$
	Thickness	h_b	2	h_p	0.66	$[mm]$
	Location	-		e_{p1}	To optimize	$[mm]$
		-		e_{p2}	169	$[mm]$
		-		e_{p3}	To optimize	$[mm]$
Material properties	Young modulus	E_b	205e9	c_{11}^S	1.20e11	$[Pa]$
				c_{12}^S	7.52e10	$[Pa]$
				c_{13}^S	7.50e10	$[Pa]$
				c_{33}^S	1.11e11	$[Pa]$
				$c_{44}^S = c_{55}^S$	2.11e10	$[Pa]$
				c_{66}^S	2.58e10	$[Pa]$
	Density	ρ_b	7850	ρ_p	7750	$[kg.m^{-3}]$
	Coupling constant	-		e_{31}	-5.35	$[C.m^{-2}]$
		-		e_{33}	15.78	$[C.m^{-2}]$
		-		e_{15}	12.29	$[C.m^{-2}]$
	Relative permittivity	-		ϵ_{11}^S	919.1	$[-]$
		-		ϵ_{33}^S	826.6	$[-]$

5.1.1.2 Placement of piezoelectric transducers to optimize modal coupling

The lateral transducers positions e_{p1} and e_{p3} must be defined to maximize the modal coupling coefficients. The variation of coupling coefficient values with transducer position follows the modal shapes shown in Figure 5.2. Due to free boundary conditions, six rigid body modes exist at low frequencies (< 1 Hz) corresponding to translation and rotation in the three spatial directions. In the following study, the frequency range of interest is below 1000 Hz and includes five bending modes and three torsional modes. In fact, the transducers must be placed on the maximum strain locations (antinodes) of the vibrations to be able to measure or act on the modes.

A parametric approach is performed and the results are shown in Figure 5.3. Because the modes are symmetrically centered, only one side result is shown, but it is representative of both sides.

The positions that maximize the coupling coefficient are not the same for all modes.The coupling coefficient for the seventh mode is greater than 0.1 regardless of the transducer position. A compromise is made between the eighth and tenth modes to maximize their coupling coefficients, so $e_{p1} = 0.07$ [m] and $e_{p3} = 0.269$ [m]. The dimensions of the designed beam are given in Table 5.2.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

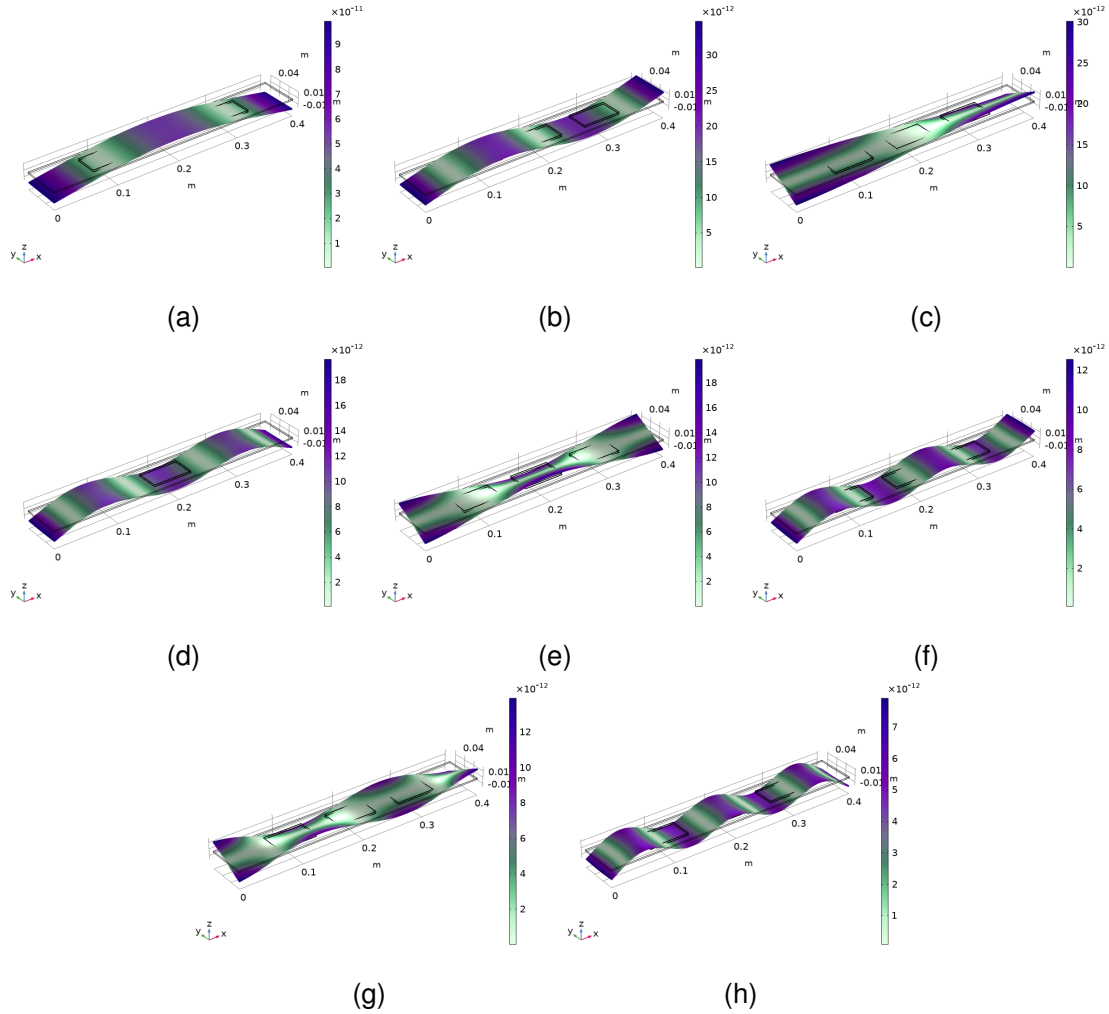


Figure 5.2: Mode shape visualization: (a) Mode 7 at 72 Hz (first bending mode), (b) Mode 8 at 194 Hz (second bending mode), (c) Mode 9 at 299 Hz (first torsion mode), (d) Mode 10 at 380 Hz (third bending mode), (e) Mode 11 at 600 Hz (second torsion mode), (f) Mode 12 at 603 Hz (fourth bending mode), (g) Mode 13 at 903 Hz (third torsion mode), (h) Mode 14 at 915 Hz (fifth bending mode).

Table 5.2: Setup dimensions

	Dimensions			Materials
	Name	Value	Unit	
Beam	L_b	400	[mm]	Steel
	b_b	60	[mm]	
	h_b	2	[mm]	
Piezoelectric Patch	L_p	61	[mm]	PI DuraAct Patch Transducer P-876.A15
	b_p	35	[mm]	
	h_p	1	[mm]	
	e_{p1}	70	[mm]	
	e_{p2}	169	[mm]	
	e_{p3}	269	[mm]	

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

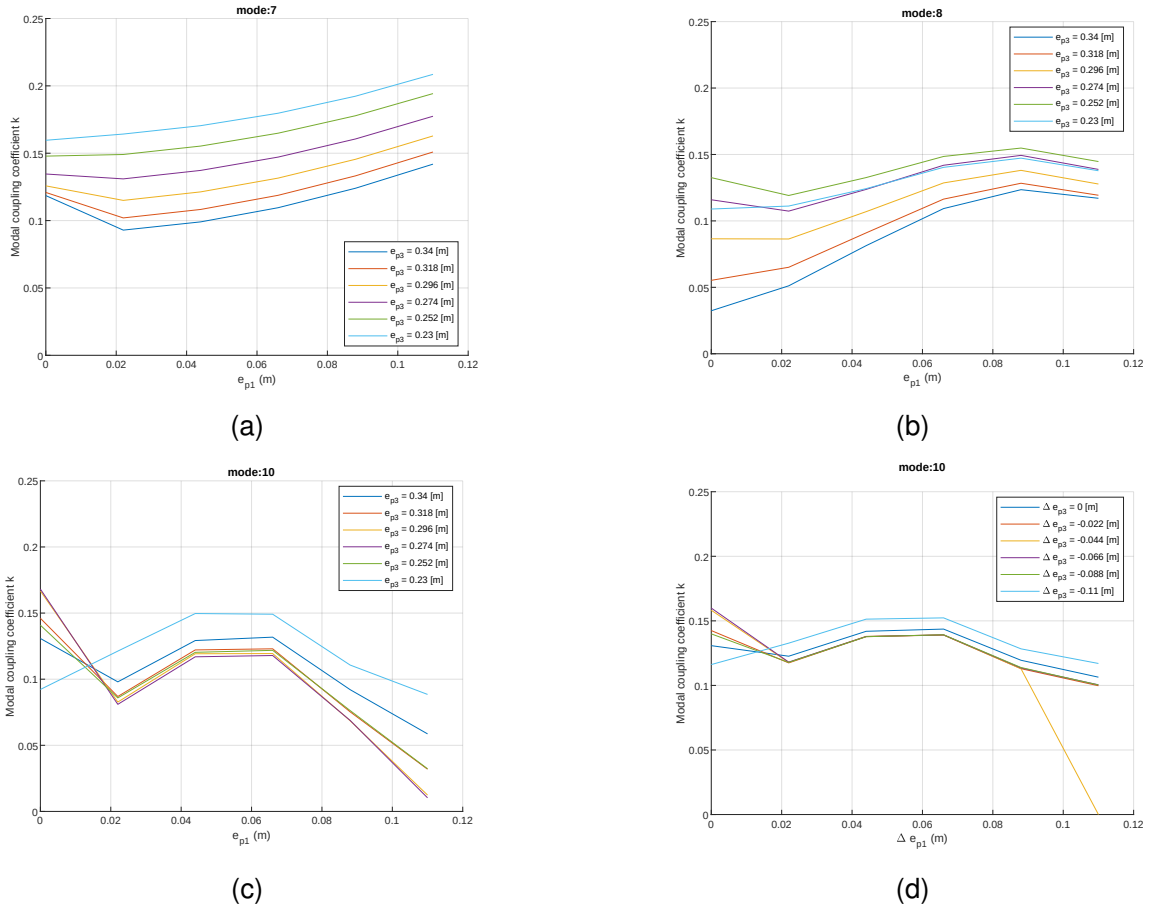


Figure 5.3: Modal coupling coefficient variations with colocalized transducer positions e_{p1} and e_{p3}: (a) mode 7, (b) mode 8, (c) mode 10, and (d) mode 10 with non-colocalized transducers where $e_{p1} = 0$ and $e_{p3} = 0.34$.

5.1.2 Experimental implementation

A test bench is built from the previous design. The bonding protocol used for the piezoelectric transducers is the same as the one presented in Chapter 3, Part 1.2. The structure properties and instrumentation are developed in this section.

5.1.2.1 Test bench setup

To satisfy the free boundary condition, the beam is suspended with white parachute rubber bands at its four corners as shown in Figure 6.1. The main support is not shown in the picture and the electrical wires have been hidden for clarity. Six transducers are added to the beam in pairs, with the top three piezoelectric transducers acting as actuators (Pa_1, Pa_2 and Pa_3) and the bottom three acting as sensors (Ps_1, Ps_2 and Ps_3).

The data acquisition setup implemented on the piezoelectric beam is shown in Figure 6.2. Here, x_1, x_2 and x_3 are the electric voltages measured at the piezoelectric transducers Ps_1, Ps_2 and Ps_3, respectively, and v_1, v_2 and v_3 are the electric voltages applied to the piezoelectric transducers Pa_1, Pa_2 and Pa_3, respectively. All devices used in the following experiments are listed in Table 5.3.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés



Figure 5.4: Experimental setup: beam with piezoelectric transducers (top view).

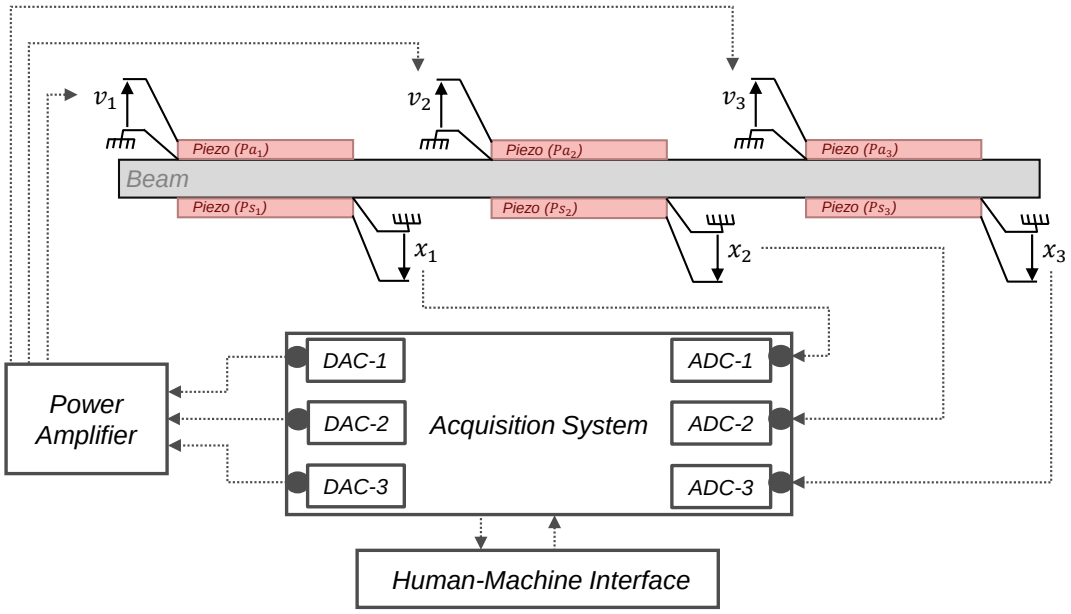


Figure 5.5: Experimental setup: data acquisition systems

Table 5.3: Experimental devices

Device	Brand	Reference	Sensitivity	Measuring Range
Acquisition System	DSPACE	DS1202	-	+/-50V
HMI	DSPACE	ControlDesk	-	-
Piezo transducers	PI	P-876.A15	0.64 µm/m/V	-250 to 1000 V
Power Amplifier	PI	DuraAct Piezo Amplifier	$<$10 mV	Input Voltage : -4 to 10 V Outpur Voltage : -100 to 250 V

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

5.1.2.2 System identification for a model building

As in Chapter 3, an experimental model is built from the measurements. The input of the identification process is a band-limited noise $v(t)$. The noise power sent to the amplifier is equal to 0.1 with a sampling frequency $F_s = 10$ kHz, such as:

$$v(t) = Rnd(t) \tag{5.1}$$

Considering the MIMO system, it is applied to the actuators one by one, while the others are short-circuited. Figure 5.6 shows measurements made with successive actuation for some selected transfer functions. Appendix C provides all measured transfer functions of the system. Note that the measurement can be performed by applying uncorrelated input noise to v_1, v_2, and v_3 simultaneously. Both methods have been tested and give similar results due to the superposition principle. Using the successive actuation over simultaneous method helps to use the coherence indicator. Even with uncorrelated noise, coherence is severely degraded with multiple actuation because other actuators act like external disturbances to the input-output transfer function under consideration. This phenomenon is not visible with the chosen method since unused actuators are short-circuited.

The measured signals x_1, x_2 and x_3 are used to identify nine transfer functions $G_{i,j}$, where $i = \{1;2;3\}$ and $j = \{1;2;3\}$, defined by:

$$\begin{bmatrix} X_1(s) \\ X_2(s) \\ X_3(s) \end{bmatrix} = \begin{bmatrix} G_{1,1}(s) & G_{2,1}(s) & G_{3,1}(s) \\ G_{1,2}(s) & G_{2,2}(s) & G_{3,2}(s) \\ G_{1,3}(s) & G_{2,3}(s) & G_{3,3}(s) \end{bmatrix} \begin{bmatrix} V_1(s) \\ V_2(s) \\ V_3(s) \end{bmatrix} \tag{5.2}$$

Where $X_1(s)$, $X_2(s)$, $X_3(s)$ and $V(s)$ are the Laplace transforms of x_1, x_2, x_3 and v respectively.

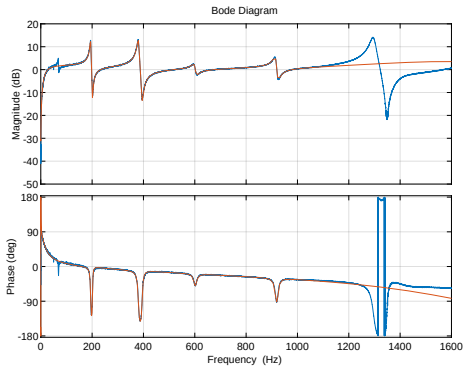
Pole and zero estimation is used to build the experimental model from the measured frequency response functions between 10 and 1000 Hz, such as:

$$G_{i=\{1;2;3\},j=\{1;2;3\}}(s) = \frac{\kappa}{s^{n_0}} \frac{\Pi_{i_z=1}^{n_z}(1 - z_{i_z}^{-1}s)}{\Pi_{i_p=1}^{n_p}(1 - p_{i_p}^{-1}s)} \tag{5.3}$$

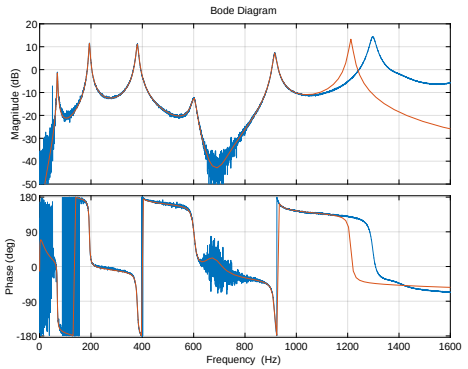
Where $\kappa \in \mathbb{R}$ is a constant gain, $n_0 \in \mathbb{Z}$ is the number of null poles (integrator if $n_0 > 0$, derivative if $n_0 < 0$), $z_i \in \mathbb{C}$ zeros and $p_i \in \mathbb{C}$ poles of the system, $n_z \in \mathbb{N}$ the number of zeros and $n_p \in \mathbb{N}$ the number of poles, $G_{i,j}$ must be proper ($n_p + n_0 \geq n_z$).

Numerical transfer functions have 13 poles and 12 zeros, except for $G_{2,1}$ which has 14 poles and 13 zeros as needed to estimate the zeros between 450 and 700 Hz. The number of poles and zeros is not a multiple of the number of modes, but they are used to introduce phase shifting at low frequencies. The first mode has a low dynamic considering the transfer functions $G_{1,1}$ (Figure 5.6a) and $G_{3,3}$ (Figure 5.6f) and is not estimated in the model. A zero amplitude measured at 0 Hz is explained by the use of piezoelectric sensors which do not measure the continuous component. A significant drop in coherence between non-collocated actuators and sensors occurs between 600 and 800 Hz (Figures 5.6d and 5.6g). In this frequency range, the model estimates zeros that are unreliable. Higher modes exist above the frequency of interest and may introduce control instabilities not considered in the model. Outside the frequency range, the sixth mode is estimated

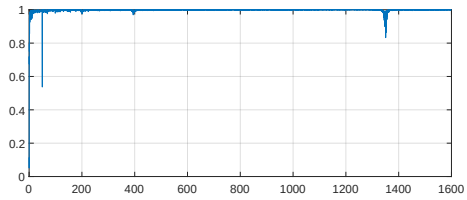
Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés



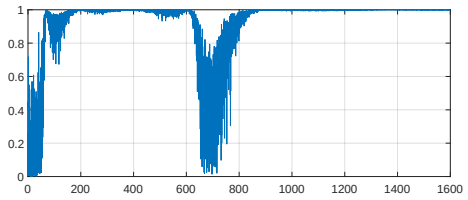
(a) $G_{1,1}$: From Pa_1 to Ps_1



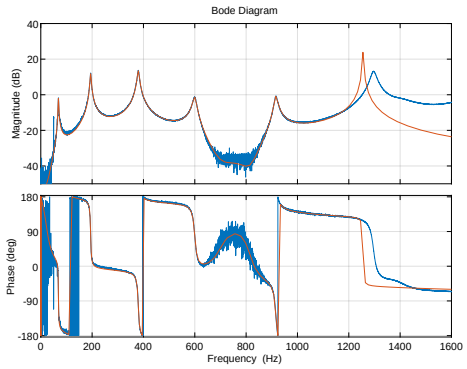
(b) $G_{3,1}$: From Pa_3 to Ps_1



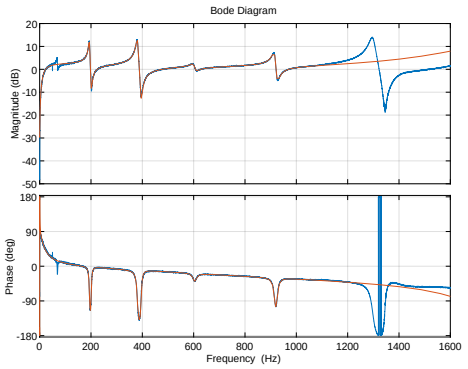
(c) Coherence: From Pa_1 to Ps_1



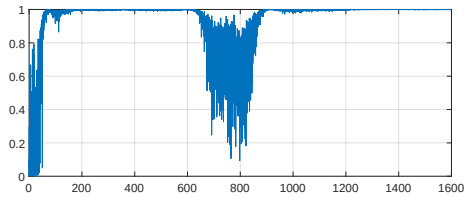
(d) Coherence: From Pa_3 to Ps_1



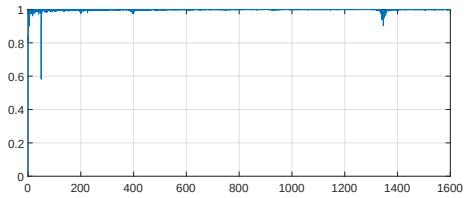
(e) $G_{1,3}$: From Pa_1 to Ps_3



(f) $G_{3,3}$: From Pa_3 to Ps_3



(g) Coherence: From Pa_1 to Ps_3



(h) Coherence: From Pa_3 to Ps_3

Figure 5.6: Bode and coherence plot with transfert function estimated from measurements (blue line) and the identified model (red line) between 0-1500 $[Hz]$

in $G_{1,3}$ (Figure 5.6b) and $G_{3,1}$ (Figure 5.6e). This allows phase fitting until 1000 Hz.

The model also fits in the time domain as shown in Figure 5.7, where the system response to a band-limited noise between 0 and 1000 Hz applied to Pa_2. Table 5.4 shows metrics that quantify the differences between the measurements and the model. The model fits reality, with an absolute error over the frequency or time domain indicator estimated from 0.2% to 7%. These differences can be explained by the non-estimation

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

of the first frequency in the $G1,1$ transfer function and the amplitude estimation of the dynamics due to the post-experiment processing of the data.

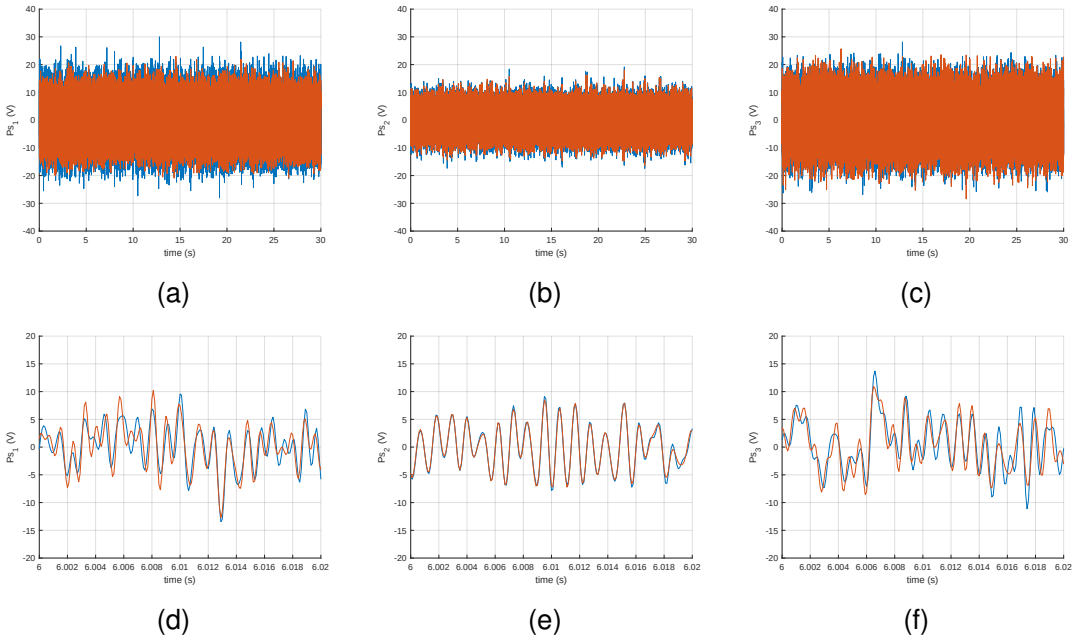


Figure 5.7: Time measurements from the experimental setup (blue line) and identified model (red dashed line) during 30 sec (a-b-c) and between seconds 6.0 and 6.02 (c-d-e).

Table 5.4: Model error quantification

Metrics	Measurements	Model	Absolute error
$\|G_{1,1}\|_\infty$	4.53	4.44	0.02
$\|G_{1,2}\|_\infty$	3.44	3.35	0.03
$\|G_{1,3}\|_\infty$	4.85	4.72	0.03
$\|G_{2,1}\|_\infty$	3.91	3.87	0.01
$\|G_{2,2}\|_\infty$	5.36	5.27	0.02
$\|G_{2,3}\|_\infty$	4.39	4.26	0.03
$\|G_{3,1}\|_\infty$	3.81	3.72	0.02
$\|G_{3,2}\|_\infty$	3.39	3.38	0.002
$\|G_{3,3}\|_\infty$	4.47	4.39	0.02
$RMS(Ps_1)$	5.15	4.79	0.07
$RMS(Ps_2)$	3.67	3.54	0.03
$RMS(Ps_3)$	5.26	5.18	0.02

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

5.1.3 Implementing control: modeling and experimental comparison

According to the structural response, the most influential bending mode is the third at 380 Hz. The control must be designed to damp this mode. Many active algorithms are efficient for vibration control, as seen in Chapter 2, Section 1.2.1. Positive position feedback control laws are suitable for vibration control of a particular mode and have the property of not introducing instabilities in higher frequency modes. It is useful in this context because high frequency modes are not considered in the model.

As in Chapter 3, Part 2.1, a feedback control law $C(s)$ is implemented on the structure defined by the system $G(s)$, as shown in the block diagram Figure 5.8. Here $r = 0$ is the target signal, e the error, $v_d = [v_{d1} v_{d2} v_{d3}]^T$ the disturbances, $v_c = [v_{c1} v_{c2} v_{c3}]^T$ the control signals, $v = [v_1 v_2 v_3]^T$ the input drive signals of the system $G(s)$ from equation (5.2) and $x = [x_1 x_2 x_3]^T$ the respective system outputs.

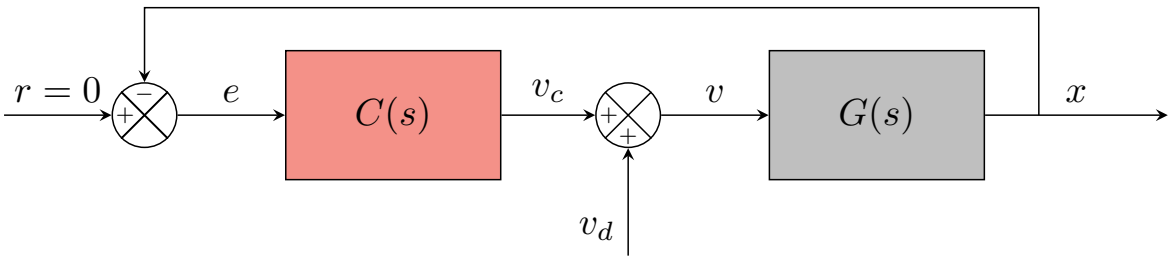


Figure 5.8: Control loop block diagram

One channel v is used to apply both the input disturbance and the control signals to the system. The signal $v_d(t)$ is a random noise $Rnd(t)$ with a sampling frequency F_s in Hz and a second order filter with a cutoff frequency $F_c = 1000$ Hz, such as:

$$\begin{bmatrix} v_{d1}(t) \\ v_{d2}(t) \\ v_{d3}(t) \end{bmatrix} = \begin{bmatrix} 0 \\ Rnd(t) * \mathcal{L}^{-1}\{\left(\frac{2\pi F_c}{s+2\pi F_c}\right)^2\} \\ 0 \end{bmatrix} \tag{5.4}$$

Where $\mathcal{L}^{-1}$ is the inverse Laplace transform and $*$ being the temporal convolution.

The feedback control laws are implemented between two collocated piezoelectric elements of the beam. Three PPF controllers are used such as:

$$C(s) = \begin{bmatrix} C_1(s) \\ C_2(s) \\ C_3(s) \end{bmatrix} = \begin{bmatrix} -Kc_1 \frac{1}{s^2+2\xi_1\omega_1 s+\omega_1^2} \\ -Kc_2 \frac{1}{s^2+2\xi_2\omega_2 s+\omega_2^2} \\ -Kc_3 \frac{1}{s^2+2\xi_3\omega_3 s+\omega_3^2} \end{bmatrix} \tag{5.5}$$

Where s is the Laplace variable of the transfer functions, Kc_i is the gain, ω_i is the filter frequency adapted to the target mode in $[rad.sec^{-1}]$ and ξ_i is the damping ratio of the control loop $i = \{1; 2; 3\}$.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

As can be seen in Figure 5.9, different PPF controllers applied to the structure have a local effect on the system response with a decrease in magnitude at 380 Hz. Tested controller parameters are referenced in Table 5.5. Measurements made on the experimental setup are compared with the model. The differences are noticeable in magnitude and phase, and are estimated to range from 10% for $G_{2,2}$ to 60% for $G_{2,1}$, as seen in Table 5.5, where the absolute error based on H_∞ norm is defined as:

$$ABS_E(|G_{2,i}|_\infty) = \frac{||Gc_{2,i}^{mod}|_\infty - |Gc_{2,i}^{mes}|_\infty|}{|Gc_{2,i}^{mes}|_\infty} \tag{5.6}$$

Where $Gc_{2,i}^{mod}$ is the modeled and $Gc_{2,i}^{mes}$ is the measured transfer function with control and $i = \{1, 2, 3\}$.

Table 5.5: Model error quantification with for PPF feedback control ($\xi = 0.23$ and $\omega = 2.39E03$)

Gain (K_c)	$ABS_E(\|G_{2,1}\|_\infty)$	$ABS_E(\|G_{2,2}\|_\infty)$	$ABS_E(\|G_{2,3}\|_\infty)$
0.05E6	0.13	0.09	0.19
0.1E6	0.27	0.10	0.21
0.3E6	0.53	0.11	0.18
0.5E6	0.55	0.37	0.20
0.8E6	0.56	0.15	0.11
2E6	0.59	0.21	0.01

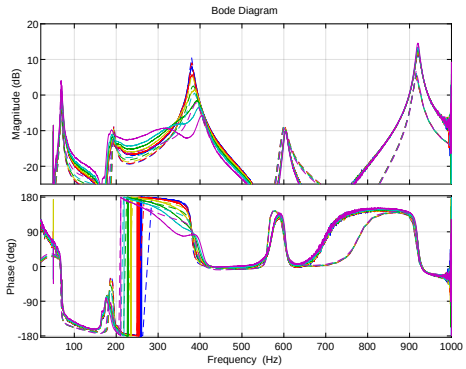
The main difference between the model and the experiment can be explained as follow:

- The effects of the controllers on magnitude and phase are underestimated by the model as observed with a phase error of at least 20 degrees on the Figures 5.9 (a-c-d). This explains why the peak doubling phenomenon induced by the PPF controller is not observable for the tested gain values.
- Higher mode dynamics are underestimated in the model, especially in $G_{2,1}$ (Figure 5.9a) where the largest error is computed. The second difference can be observed close to 380 Hz on the Figures 5.9 (b-d-f).

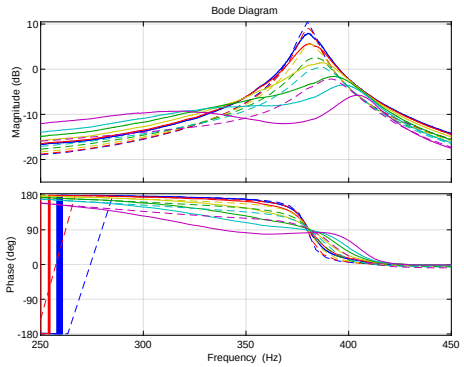
Both differences introduce estimation errors for instabilities. As the controller gain increases, the amplitudes of the first and fifth bending modes at 70 and 920 Hz increase, approaching spillover and leading to system instability. According to the simulations, the first bending mode is more likely to affect the gain margin and the fifth mode the phase margin. These illustrate that even with a suitable model, errors can be introduced by the addition of a feedback loop on an experimental test bench.

To ensure a gain margin of 10dB, a reference controller is manually set from the previous results, identical for the three loops, with a gain coefficient of $K_c = -0.4e6$, $\xi = 0.23$ and $\omega = 2.39e03$, with the aim of robustness. For the optimization methods used in the following parts, the controllers for the three loops can be tuned with different parameters.

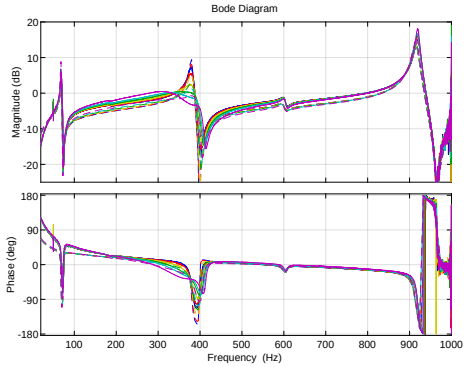
Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés



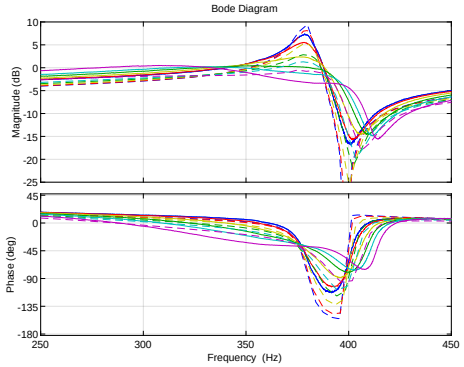
(a) From Pa_2 to Ps_1



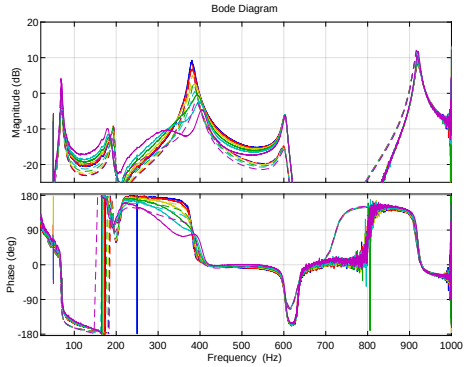
(b) From Pa_2 to Ps_1 (zoom)



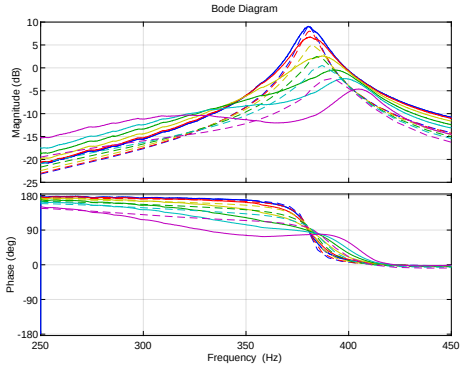
(c) From Pa_2 to Ps_2



(d) From Pa_2 to Ps_2 (zoom)



(e) From Pa_2 to Ps_3



(f) From Pa_2 to Ps_3 (zoom)

Figure 5.9: Bode plot for different controllers (Table 5.5) applied to the system: comparison of model (- -) and experimental (–) measurements where $\xi = 0.23$, $\omega = 2.39E03$, $K_c = 0.05E6$ (blue) , $K_c = 0.1E6$ (red), $K_c = 0.3E6$ (orange), $K_c = 0.5E6$ (green), $K_c = 0.8E6$ (cyan) and $K_c = 2E6$ (magenta).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

5.2 Control tuning methods

This section examines two methods for tuning PPF controller parameters. The first uses a deep reinforcement learning algorithm with the same methodology as used in Chapter 4. The second method uses a simplex algorithm defined in Chapter 2 Part 1.3.3. Both optimization methods are performed on a defined environment. The control objective is set by the previously manually tuned controller in Part 5.2.1.2. A statistical study is carried out to compare the methods, taking into account computation time, control performance and robustness to environmental variations.

5.2.1 Environment definition

As in Chapter 3, the environment is defined by the previously estimated model from equation (5.2). It consists of observations, actions and rewards. These are calculated from the metrics computed from the model.

5.2.1.1 Observations

Three observations are defined as the root mean square of the time signal measured by the sensors during 30 seconds such as:

$$S = \begin{bmatrix} S_1 \\ S_2 \\ S_3 \end{bmatrix}_{v_d} = \begin{bmatrix} rms(x_1) \\ rms(x_2) \\ rms(x_3) \end{bmatrix} \tag{5.7}$$

This metric is relevant considering the noise excitation of the beam and can vary according to the frequency range of the excitation due to the modal distribution of the beam. Note that these values are normalized between -1 and 1 to avoid having large values as NN input, which can degrade it's training. The environment does not change over time, so the agent only interacts with it once before adjusting its policy. Thus, the observation definition plays a minor role in this context, since the trained network is only adapted to one configuration of the beam and takes only one set of observations as input.

5.2.1.2 Action

Nine actions have to be tuned according to the three PPF controllers implemented on the structure. They are grouped in an action variable $A \in \mathbb{R}^{3\times3}$ such as:

$$A = \begin{bmatrix} K_{c1} & \xi_1 & \omega_1 \\ K_{c2} & \xi_2 & \omega_2 \\ K_{c3} & \xi_3 & \omega_3 \end{bmatrix} \tag{5.8}$$

The limits of the damping parameters are set by the controller definition (Chapter 2 Part 1.2.1) with $\xi_1 \in [0; 1]$. The filter frequency parameter is limited by the model frequency range with $\omega_i \in [0; 6283]$. The controller gain limits are set in a wide range of values, such as : $Kc_i \in \{-1 \times 10^0; ...; -7 \times 10^{10}\}$. This range of parameters includes both stable and unstable controllers.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

5.2.1.3 Reward

The reward definition is designed with two objectives: firstly, to minimize the vibrations of the beam at 380 Hz and, secondly, to ensure the stability of the controlled structure. It is defined using a frequency domain integration between $f_1 = 350$ and $f_2 = 388$ Hz to consider only the dynamics of the poles associated with the third bending mode at 380 Hz. It contains the transfer functions $G_{1,1}$, $G_{2,2}$ and $G_{3,3}$ in a ratio where the control is off and on, such as:

$$R_1 = 20 \times log_{10} \sum_{i=1}^{3} \frac{\int_{f_1}^{f_2} G_{i,i}^{off}\, df}{\int_{f_1}^{f_2} G_{i,i}^{on}\, df} \tag{5.9}$$

This equation defines the reward for a stable controlled system where gain margins (Gm) are above 10dB (Chapter 2, Section 2.1.2, paragraph 2.1.2.1). For systems operating within a limited stability range ($Gm < 10dB$), a stability constraint is introduced into the reward definition with the gain margin, such as:

$$R_1 = 20 \times log_{10} \sum_{i=1}^{3} \frac{\int_{f_1}^{f_2} G_{i,i}^{off}\, df}{\int_{f_1}^{f_2} G_{i,i}^{on}\, df} \times |\prod_{i=1}^{i=3} \frac{Gm_{dB}^{i}(0 < Gm_{dB}^{i} < 10)}{10}| \tag{5.10}$$

Where Gm_{dB}^{i} is the gain margin of the loop $i = \{1, 2, 3\}$ when the other two controllers are applied to the system. In case of an unstable system, the reward is defined such as:

$$R_1 = -1 \tag{5.11}$$

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

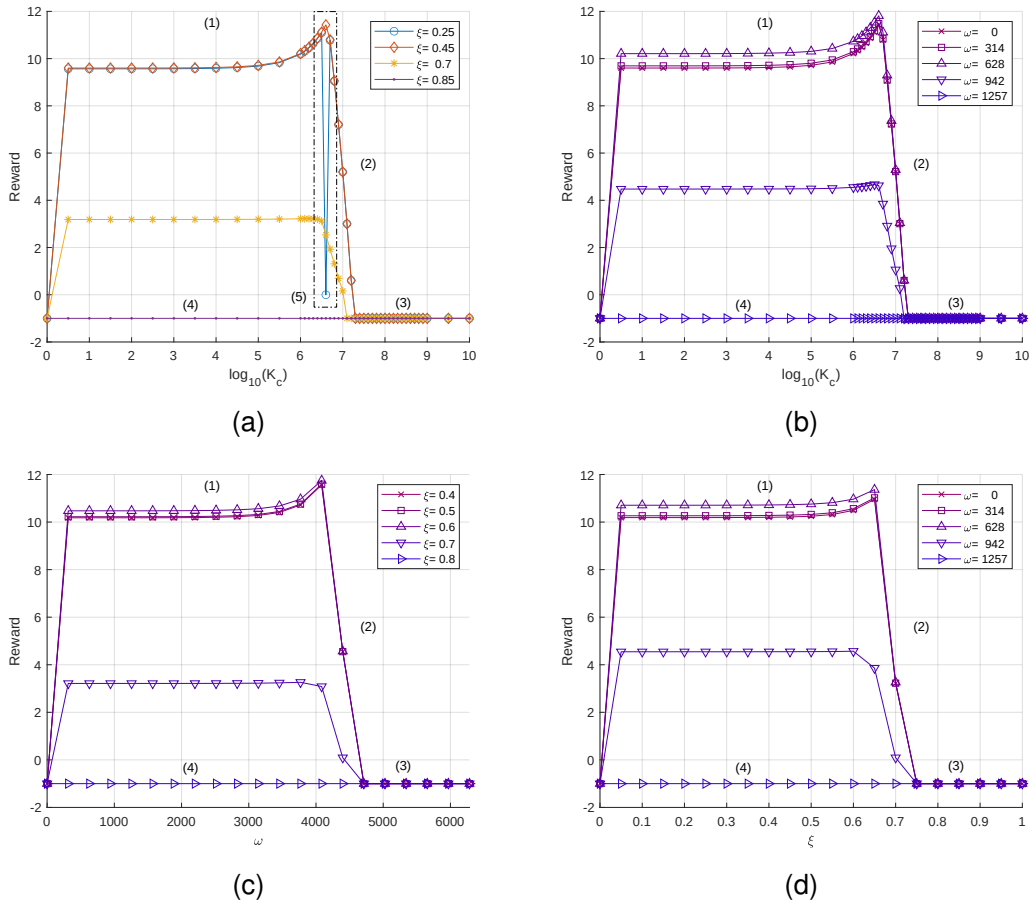


Figure 5.10: Reward R_1 variation with control law parameters set as identical for all control loops: $\omega_{1/2/3} = 380 \times 2\pi$ (a), $\xi_{1/2/3} = 0.55$ (b), $Kc_{1/2/3} = 1e6$ (c-d).

The conditional reward definition permits a transition from the stable to the unstable region into the algorithmic process, thus avoiding discontinuity in the region of the instability.

- Figure 5.10a shows the reward variation with gain K_c for different damping coefficients ξ and a fixed frequency $\omega_i = 380 \times 2\pi = 2.39E03$ (rad/s). A numerical artefacts can be observed for $\xi = 0.25$ and $log_{10}(K_c) = 6.6$ with a drop in reward (5).
- Figure 5.10b shows the reward variation with gain K_c for different frequencies ω and a fixed damping coefficient $\xi = 0.55$.
- Figure 5.10d shows the reward variation with gain ξ for different frequencies ω and a fixed gain $K_c = 1e6$.
- Figure 5.10c shows the reward variation with gain ω for different frequencies ξ and a fixed gain $K_c = 1e6$.

In these examples, the controllers are identical for all three loops to provide a 2-dimensional visual representation. The reward increases with the parameter value (K_c, ω or ξ)(1) and then gradually decreases as the gain margin decreases (2). High gain values lead to an unstable system (3). In addition, high values of frequencies ω and damping ratio ξ lead to an unstable controlled system (4).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

5.2.2 Methods comparison

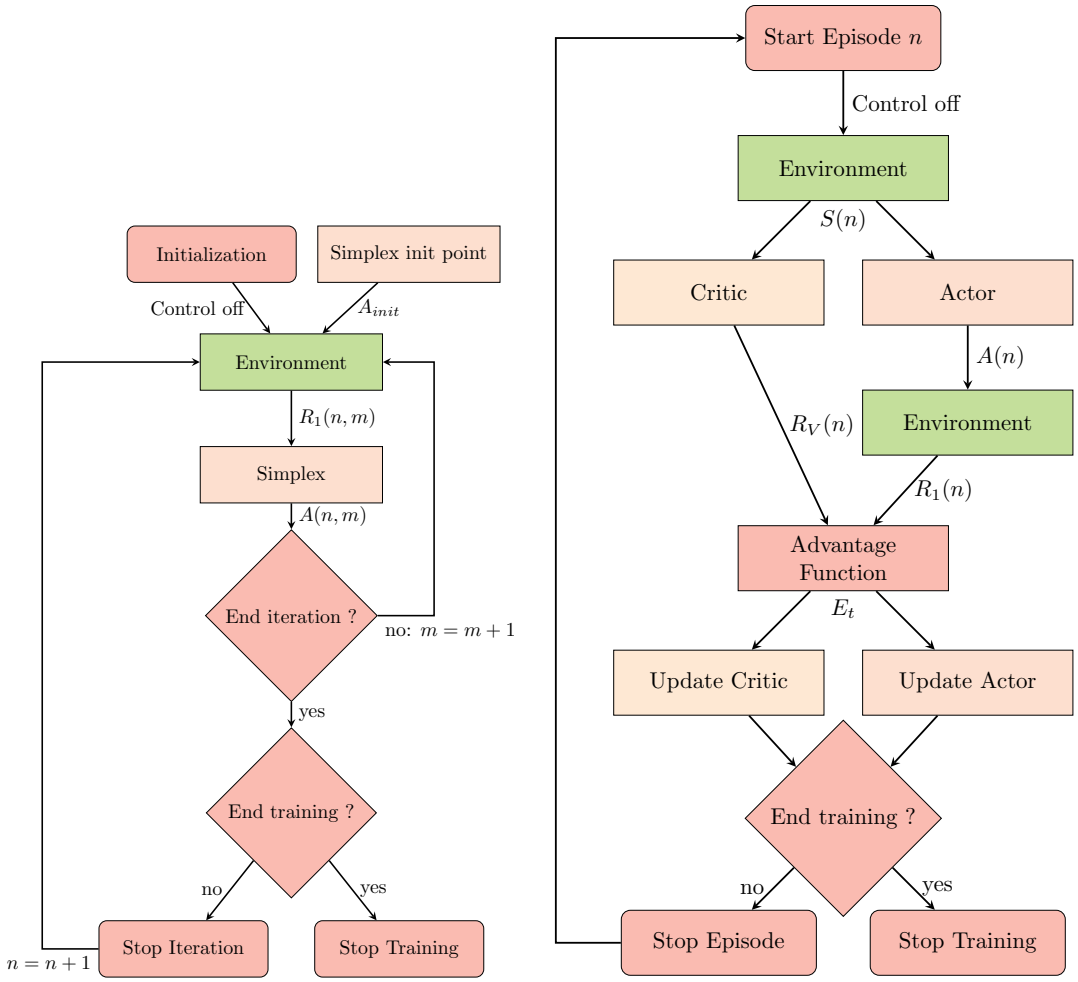
5.2.2.1 Simplex

The Simplex algorithm evolves with an objective function represented by the reward computed from the environment in this case. To consider the minimization process of the algorithm, the objective function is chosen with a negative sign of the conditional reward defined in equations (5.9), (5.10) and (5.11). Figure 5.11a shows a simplified flowchart of the algorithm in this application case. Note that the inner workings of the Simplex algorithm are detailed in Chapter 2 Part 1.3.3, and $m \in \mathbb{N}$ is the number of function evaluations in an iteration $n \in \mathbb{N}$. The algorithm starts with an initial point, defined here by a set of controller parameters A_{init}. After initializing the environment (control off), the initial controller can be applied to the structure to compute the associated reward R_1. Then the simplex optimization algorithm is executed until the termination state is reached. It can be terminated by one of two conditions: the number of solver iterations n exceeds 1800, or the tolerances on the objective function and parameters are below 1e-4. During one iteration, the algorithm can evaluate the objective function only 1800 times. The performance of the algorithm is supposed to be highly dependent on the starting point, so several must be tested.

5.2.2.2 Deep reinforcement learning

The deep reinforcement learning algorithm uses TRPO method as developed in Chapter 3 Part 3.2 with the same default parameters and the same neural network size. The agent composed of two neural networks, the actor and the critic interact multiple times with the environment to update his policy according with the reward R_1 increases. Figure 5.11b is a flowchart of the optimization algorithm main steps. The formal equations of the TRPO method are detailed in Chapter 2 Part 2.2.2. The training is divided into episodes. An episode n starts with the initialisation of the environment (control off). The initialization states are used as input to the neural network to estimate the expected reward R_V for the critic and the actions $A(n)$ (controller parameters) of the actor. The selected controller is evaluated within the specified environment to compute the corresponding episode reward $R_1(n)$. The reward function is calculated using the expected and actual rewards (R_V and R_1) in order to update the bias and weight of the actor and critic neural networks. Training can end under two conditions: the computed reward reaches a value defined by the user, or the maximum number of episodes is reached. In this example, the maximum number of episodes is set to 5000 and the maximum reward is set to 20. This value is high for this application and cannot be reached given the stability constraint in the reward definition computed from the model. This way, the training will be stopped by the number of episodes condition. The number of episodes is chosen as relatively low compared to the previous chapter, considering that the number of parameters is multiplied by three (3 parameters to tune in Chapter 3 and 9 in this chapter). The ability to converge was highlighted in Chapter 3 after a significant number of episodes. Divergences may occur depending on the NN initialization, so several must be tested.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés



(a) Simplex (b) Deep Reinforcement Learning (TRPO)

Figure 5.11: Flowcharts

5.2.2.3 Statistical results

The parameters initialization step is an input condition of both methods. The aim of this numerical experiment is to determine the success rate without prior knowledge of the optimal results using random initialization. The convergence speed and control efficiency of the two methods are thus studied using a statistical approach. 85 agents are randomly initialized with different seeds to ensure numerical randomness also in the DRL training process. The starting points used for the simplex optimization are actions A_{init} given by the initialized agents as shown in Figure 5.12.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

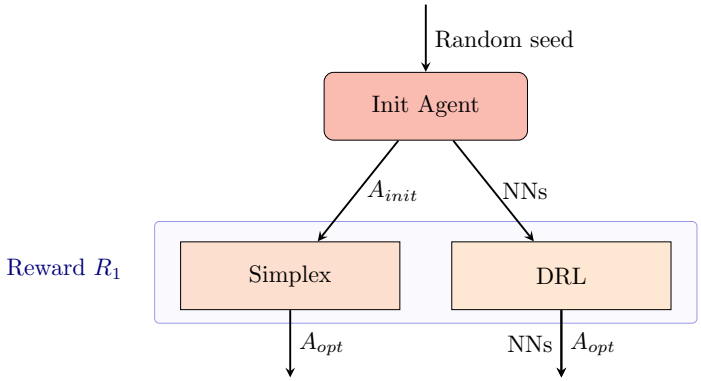


Figure 5.12: Flowchart: DRL and Simplex with a single initialization

Figure 5.13 shows the distribution of simplex and DRL methods according to the reward obtained at the end of the optimization.

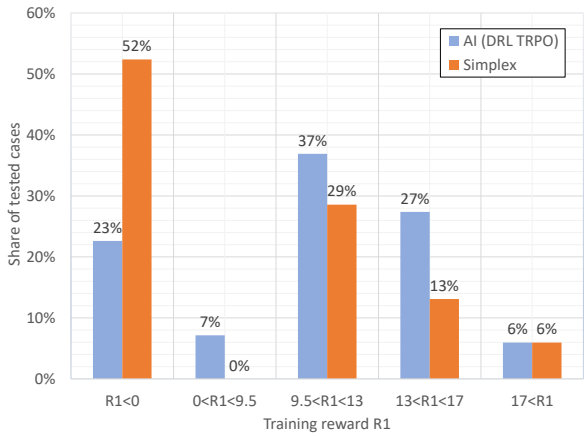


Figure 5.13: Proportion of tested cases according to the reward obtained at the end of the calculation (considering 85 interdependent cases tested)

A reward of 9.5 corresponds to the structure when control is off. For simplex, 52% of the tested cases lead to a controlled system that is worse than the uncontrolled one, compared to 30% for DRL. If the controller is effective (reward above 9.5), cases are divided into three categories:

- $9.5 < R1 < 13$: Moderate control efficiency.
- $13 < R1 < 17$: Expected control efficiency.
- $17 < R1$: Unexpected control efficiency

A reward of 13 is equivalent to the previously established reference controller efficiency. These results need to be analyzed in the light of the number of environmental assessments required to achieve the lower reward of the category available in Figure 5.14.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

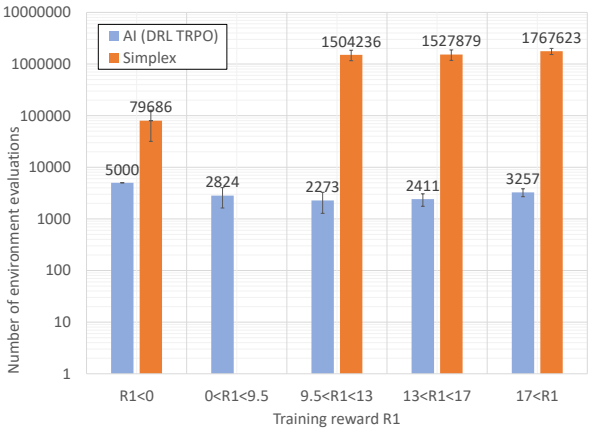


Figure 5.14: Means and standard deviations of the number of environmental evaluation shared according to the reward obtained at the end of the calculation (considering 85 inter-dependent tested cases).

These figures highlight the convergence speed of the DRL method compared to the Simplex method. For the Simplex, the number of environment evaluations can reach 3.24 million (according to the options defined for the termination conditions). For all categories, the Simplex stopped because of the tolerance conditions. It can be deduced that Simplex leads to find a local optimum as a solution. All DRL algorithms are trained on 5000 episodes, but not all episodes are necessary to reach the different categories.

Figure 5.15 shows the reward variations with epidode during DRL training, divided into the above categories. Even after 5000 episodes, only 6% of the tested agents manage to reach the assumed optimal reward value. This can lead to three main hypotheses.

- First, other agents need more episodes to reach the same reward. As can be seen in figures 5.15a and 5.15b, the reward has an increasing tendency and does not seem to have reached a ceiling.
- Second, this maximum reward value is reached with a discontinuity according to the parameters as observed in Figure 5.10a.
- Thirdly, the agent diverges especially for categories that give an unstable control system.

It has been observed that some agents estimate action values outside the defined range, increasing with episodes and leading to infinite values.

The DRL method is more efficient than the Simplex in terms of the proportion of stable cases found for a 700 times smaller number of environment evaluations. However, the DRL method consumes more computational time and electric power to run. Both methods were run on the same computing server unit (properties in Appendix F). DRL uses 8 GB of CPU in 2 hours 19 minutes and 16 seconds to run 5000 episodes, giving an indicator of 0.2 seconds per environment evaluation per GB. Simplex uses 20 GB of CPU in 1 hour 22 minutes 27 seconds for 1223606 episodes, giving an indicator of 0.0002 seconds per environment evaluation per GB. Simplex interacts with the environment 700 times more than DRL, although it is 1000 times faster. These differences in computation time are explained by the updating process of the weight and bias of the neural networks and the initialization of the environment for each episode required for the DRL method.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

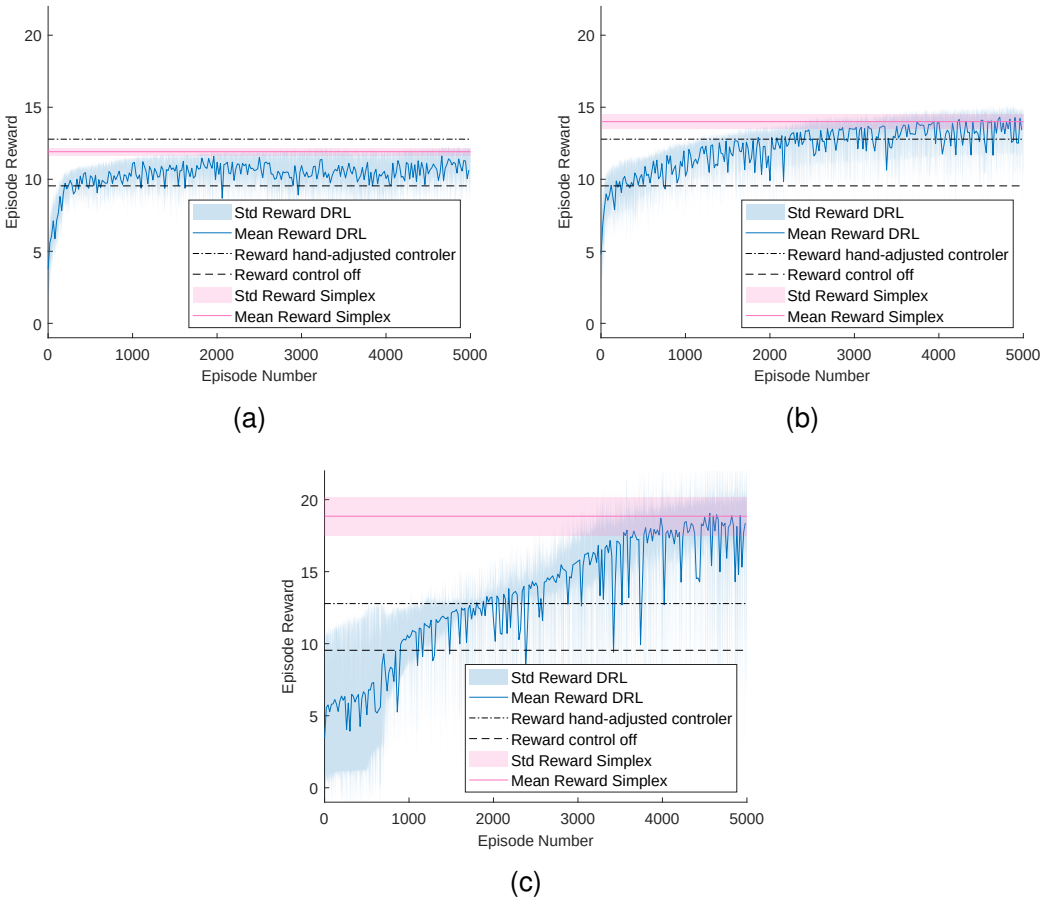


Figure 5.15: DRL reward variation with episodes grouped by final reward value achieved: between 9 and 13 (a), between 13 and 17 (b), and over 17 (c) (The figures do not represent simplex variations with an episode; it only represents the terminal state.)

5.2.3 Control performances on the model

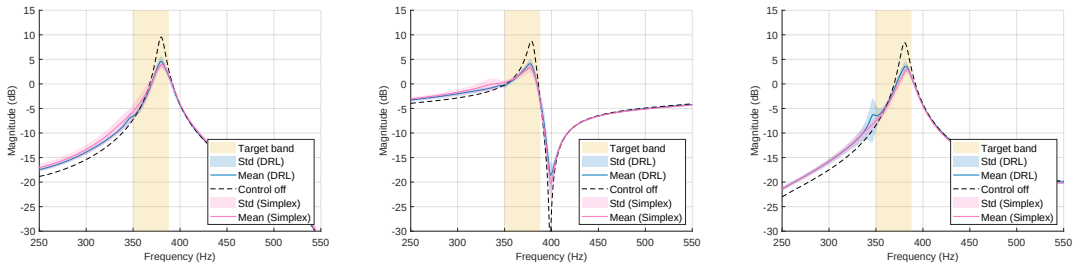
The Simplex and DRL methods resulted in the generation of tuned controllers exhibiting a range of degrees of efficiency, as quantified by the reward value. The controllers are tested on the model and their influence on the dynamics of the system is shown in Figure 5.16 and grouped according to the above range of reward categories. Appendix C provides large frequency response of the controlled numerical model.

- Figures 5.16a, 5.16b and 5.16c show moderate effects of the control law on the target mode. The gain reduction is about 5 dB for the two tested methods.

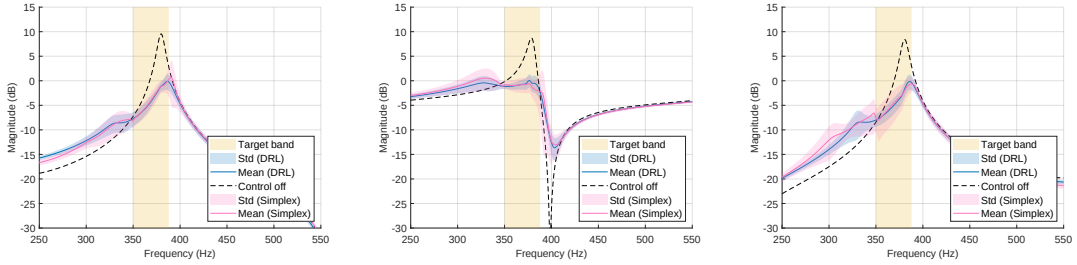
- Figures 5.16d, 5.16e and 5.16f show the expected effects of the control law on the target mode. The gain reduction is between 8 and 10 dB for DRL and between 5 and 10 dB for simplex within the target frequency band. The selected controls serve the function of a secondary tuned mass damper system. This result in an amplitude split from one side of the primary target frequency to the other.

- Figures 5.16i, 5.16h and 5.16g show aberrant effects of the control law on the target mode. The gain reduction is between 10 and 30 dB for DRL and between 10 and

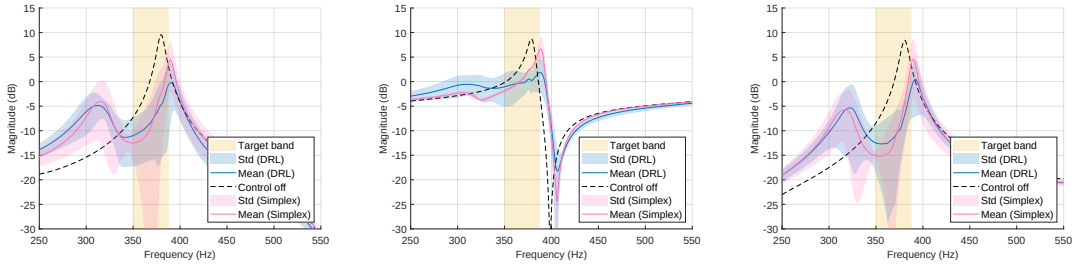
Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés



(a) From Pa_2 to Ps_1 (b) From Pa_2 to Ps_2 (c) From Pa_2 to Ps_3



(d) From Pa_2 to Ps_1 (e) From Pa_2 to Ps_2 (f) From Pa_2 to Ps_3



(g) From Pa_2 to Ps_1 (h) From Pa_2 to Ps_2 (i) From Pa_2 to Ps_3

Figure 5.16: Bode plot for different controllers applied to the model: final reward between 9 and 13 (a-b-c), final reward between 13 and 17 (d-e-f) and final reward over 17 (g-h-i).

40 dB for Simplex at 380 Hz. This controller performance is unexpected. Further investigation reveals that these chosen controllers are unstable when applied to the test-bench. In fact, the phase margins are all negative. These controllers tuning exploit the limits of the reward definition: the performances exceed the gain margin stability constraint penalty and the numerical computation was not able to identify the system instabilities.

Also for most of the above tuned controllers it is not possible to perform measurements on the experimental setup due to unstable tuned controllers. Experimental uncertainties have not been considered in the modeling and the reward definition is unable to compensate for them.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

5.3 Robustness of tuning methods to reward variation

During the lifetime of a structure, it is subject to variations related to its operation (e.g. fatigue or damage) or to external environmental conditions (temperature, humidity...). The use of an adaptive method would allow the control parameters to be effectively adjusted to take account of these variations. This section explores the properties of the DRL to adapt to changes in the environment by changing the reward definition. Previous results are used as a baseline so that only the tests that provide significant results are reused, excluding unstable controlled model cases.

5.3.1 Target frequency bandwidth expansion

The numerical experiment is conducted using the most optimal training data previously obtained. This corresponds to the agent that reaches the maximum final reward above 20. The aim is to observe the capacity of the DRL to adapt to the control objective. The target frequency range is expanded by introducing new rewards derived from equations (Eq. 5.9) to (Eq. 5.11), as follows:

- For reward R_1: $f_1 = 350$ Hz and $f_2 = 388$ Hz
- For reward R_2: $f_1 = 300$ Hz and $f_2 = 388$ Hz
- For reward R_3: $f_1 = 220$ Hz and $f_2 = 388$ Hz

Figure 5.17 illustrates the strategy of continuing training with reward variation from R_1 to R_3:

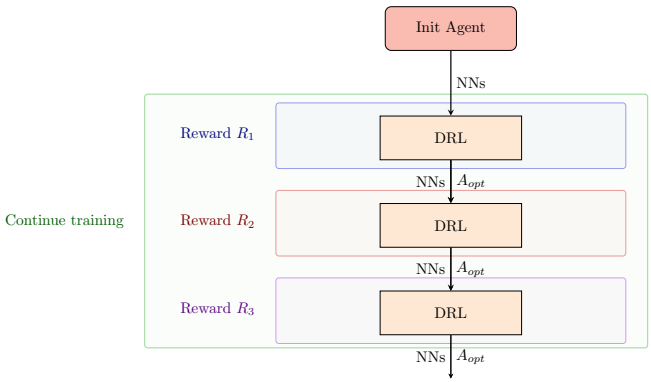


Figure 5.17: Flowchart: continuing the training strategy

- The initial agent is already trained during 5000 episodes with the goal of maximizing the reward R_1.
- Then the trained agent is again trained during 5000 episodes with the goal of maximizing the reward R_2.
- Finally, the same agent trained on reward R_2 is trained on reward R_3.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Figure 5.18 shows reward variations with episodes. The changes in the reward definition are indicated by the background colors and the zone indicators (1), (2) and (3). The following elements can be observed:

- Transitions between zone (1), (2) and (3): The range of reward values, and therefore the maximum values that can be achieved, decreases as the range of target frequencies increases.

- Zone (2) and (3): Reward R_2 converges to the upper bound in less than 500 episodes, while reward R_3 converges in more than 2300 episodes.

These observations highlight the relationship between the adaptation time (number of episodes) and the variation in the environment to reach the updated optimum. Further calculations are required to establish a relationship between the variation in the environment and the number of episodes that DRL needs to adapt to it.

The controllers tuned according to each reward definition are tested on the transfer functions model. Figures 5.19a, 5.19c and 5.19e show the effect of the controllers on the dynamics of the structure, focusing on the transfer functions of the collocated transducers. Figures 5.19b, 5.19d and 5.19f show the controllers functions.

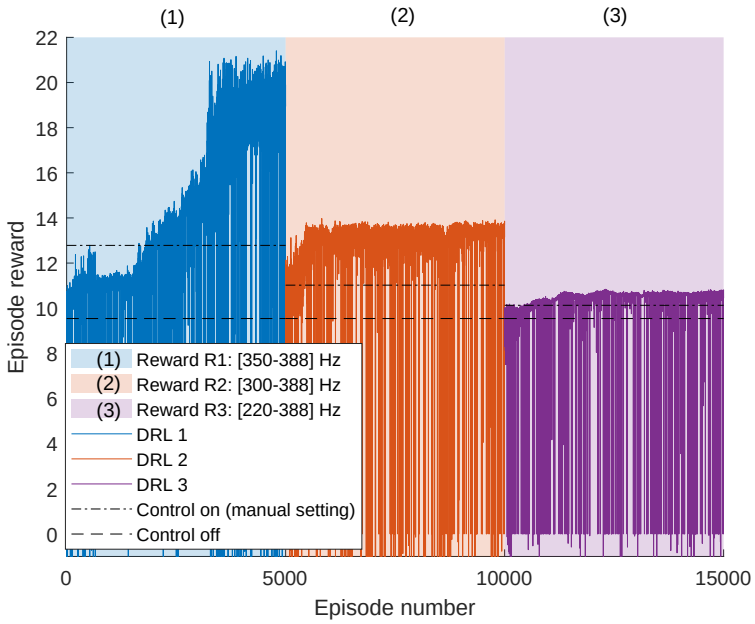


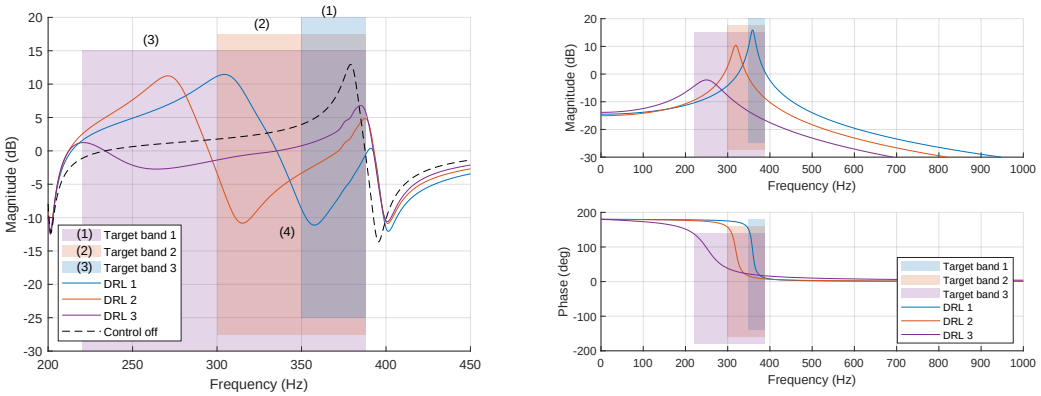
Figure 5.18: Reward variation with episodes using the continuing training strategy. (The manual settings are defined in part 5.1.3

The agent has tuned a set of controllers with different effects on the transfer functions:

- Figure 5.19a: All controllers have a higher influence on the transfer function from Pa_1 to Ps_1.

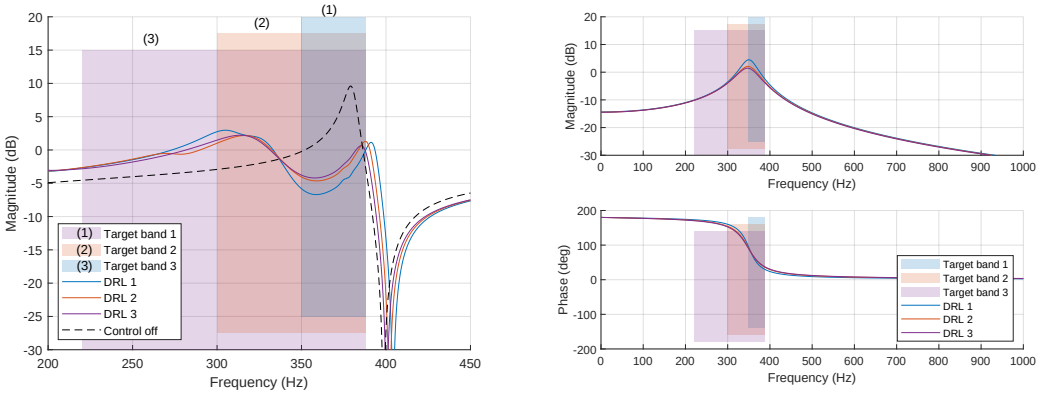
- Figure 5.19a-(4): Considering the reward R_1, the optimal controller introduces a zero in the target frequency band. Enlarging the band moves the zero to lower frequencies. Changing the reward has little effect on the other two collocated transfer functions.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés



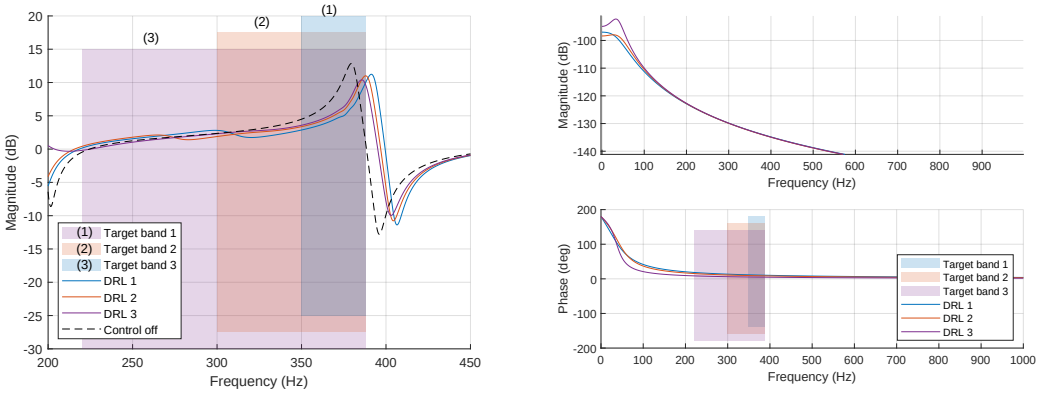
(a) From Pa1 to Ps1

(b) Controller loop 1



(c) From Pa2 to Ps2

(d) Controller loop 2



(e) From Pa3 to Ps3

(f) Controller loop 3

Figure 5.19: Bode plot for three tuned controllers applied to the model with increasing target band frequency range.

- Figure 5.19b: The controller used on the first loop has a low damping coefficient that increases with the range of the target frequency.
- Figure 5.19d: The controllers used on the second loop look the same regardless of the reward definition.
- Figure 5.19f: The third controller uses low gains and has no effect on the system with zero phase in the target frequency band.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

As a consequence of the global optimisation of the target frequency range, the reward leads to an unrealistic case where the focus is on the optimisation of a single transfer function in order to achieve a balance between high controller performance and the numerical stability of the system. Nevertheless, this example serves to illustrate the capacity of the DRL algorithm to adapt to variations in the surrounding environment. A query has been raised regarding the rate of adaptability when training is conducted directly from R1 to R3 without R2. It would be beneficial to ascertain whether this process is slower or faster.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

5.3.2 Stability criteria improvement

This study compares the performance of DRL and Simplex in adapting to reward constraints and highlights the impact of using a pretraining on the resulting performance.

The reward gain margin is not enough to ensure stability within the test bench environment. To address this issue, the phase margin Pm^i_{deg}, for each control loop $i = \{1; 2; 3\}$ in degree, is used in the reward definition for equations (5.9), (5.10) and (5.11), such as:

- Stable case: $R_4 = 20 \times log_{10} \sum_{i=1}^{3} \frac{\int_{f_1}^{f_2} G^{off}_{i,i}\, df}{\int_{f_1}^{f_2} G^{on}_{i,i}\, df}$
- Limit of stability case:
  $R_4 = 20 \times log_{10}(\sum_{i=1}^{3} \frac{\int_{f_1}^{f_2} G^{off}_{i,i}\, df}{\int_{f_1}^{f_2} G^{on}_{i,i}\, df} \times |\prod_{i=1}^{i=3}(\frac{Gm^i_{dB}(0<Gm^i_{dB}<10)}{10})| \times |\prod_{i=1}^{i=3}(\frac{Pm^i_{deg}(0<Pm^i_{deg}<180)}{180})|$
- Unstable case : $R_4 = -1$

Figure 5.20 shows the new defined reward variations with controllers parameters. Including the phase in the definition constrains the reward value when a controller is selected near the instability region. As can be seen on Figure 5.20a, the reward drops off faster than with the previous definition in Figure 5.10a. This includes more penalty in the reward for being close to the instability range.

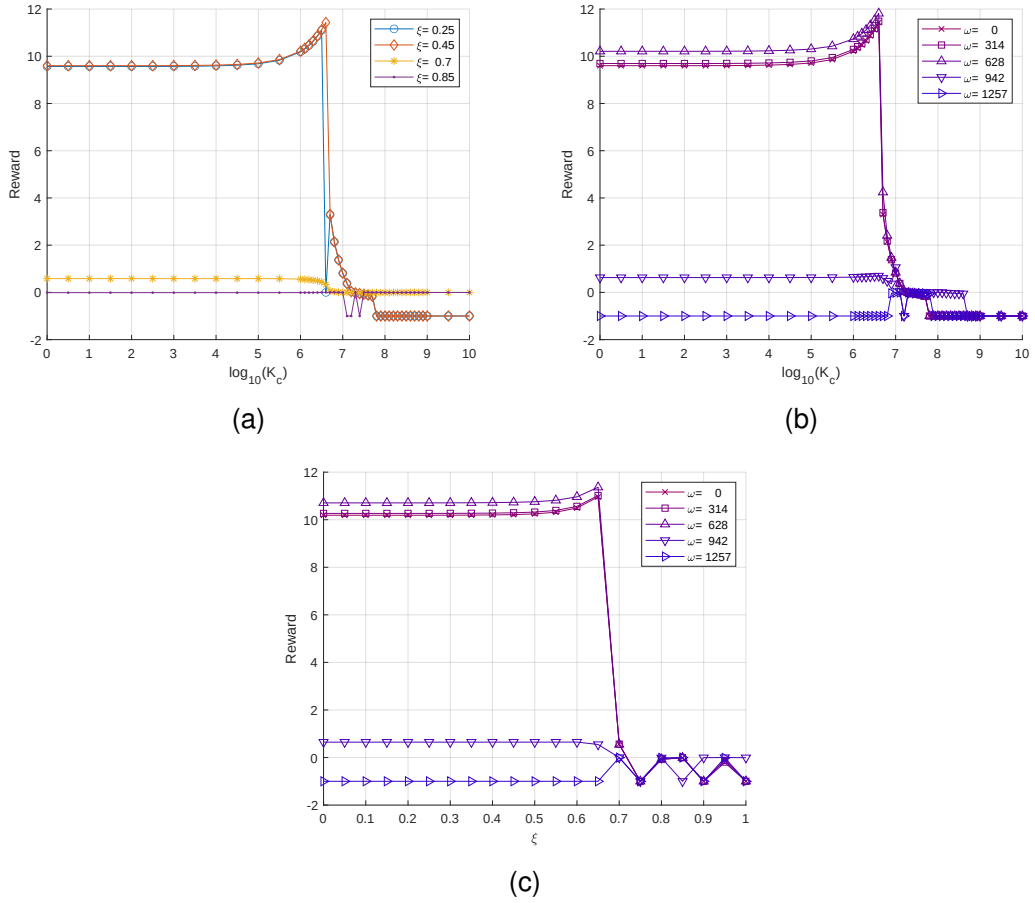


Figure 5.20: Reward R_4 variation with control law parameters set as identical for all control loops: $\omega_{1/2/3} = 380 \times 2\pi$ (a), $\xi_{1/2/3} = 0.55$ (b), $Kc_{1/2/3} = 1e6$ (c).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Using the previous results of Part 5.2.2, only the NNs coming from the 20 best trainings are selected. This choice gives the DRL method an advantage over the simplex method, as the best initialization for DRL is not the same as the best initialization for the simplex method. Nevertheless, the aim is to facilitate a comparison between the two methods in terms of their convergence rate and capacity to adapt to changes while starting from the same point.To adjust the controller according to the new reward definition R_4, DRL and simplex methods are investigated as introduced in Figure 5.21:

- The first strategy uses the initialized agents and restarts training with the goal of maximizing R_4.
- The second approach uses the final agent trained on R_1 as the initial state to continue training on R_4.

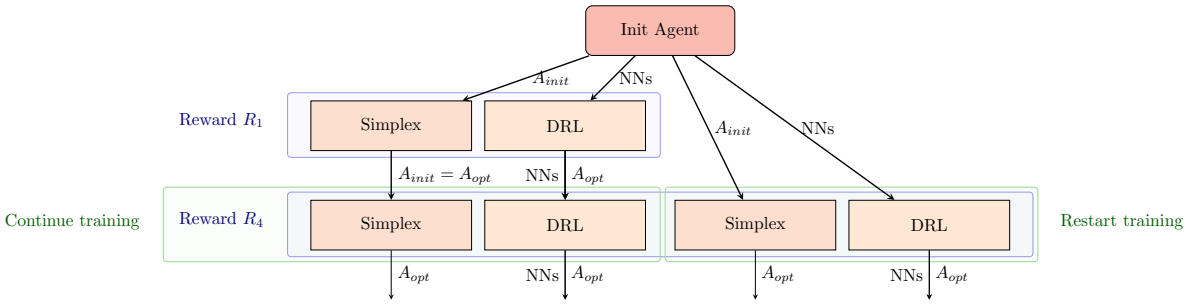


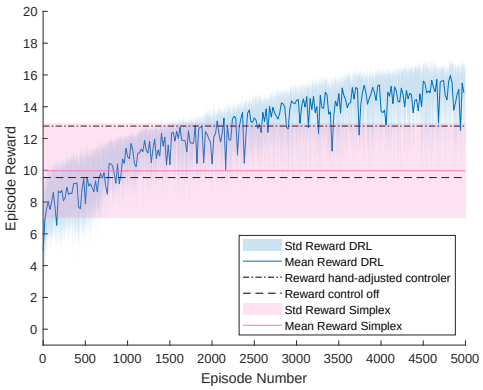
Figure 5.21: Flowchart: restarting or continuing the training strategy

Figure 5.22 shows the different training strategies and the control efficiency on the modeled structure.

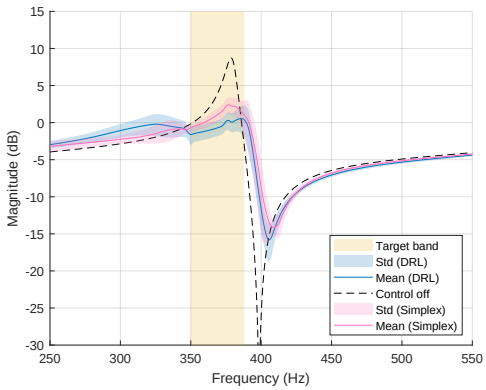
- Figure 5.22c and 5.22e: A comparison of the training after 5000 episodes reveals that the restart strategy fails to reach the target control efficiency, in contrast to the continuous strategy.
- Figure 5.22a: The mean reward is biased by the high reward values, which can reach over 17 with reward definition R_1 even for unstable controlled systems, which is not the case for R_4.
- Figures 5.22b and 5.22f illustrate similar controller performance on the system. The influence of the stability criteria on the system response occurs at low frequency (71 Hz) with a decrease in the first mode amplitude gain with R_4 optimization.
- Figures 5.22b and 5.22f: Simplex tuning methods are still less efficient than DRL with a difference of 2 dB on the magnitude at the target frequency mode (380 Hz).

The point here is to show the efficiency of the method in achieving a reference target, not to exceed it. This would require more episodes if better performances were achievable.

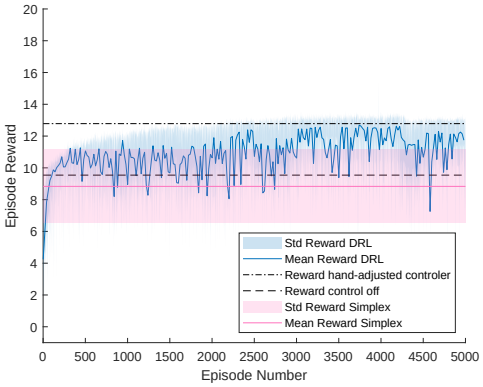
Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés



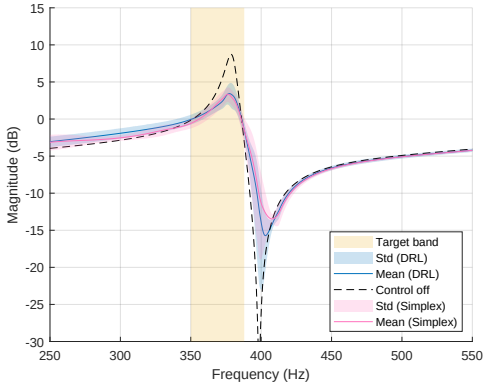
(a) Initial training with R_1



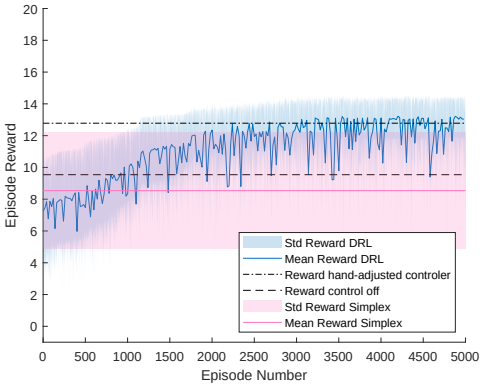
(b) From Pa_2 to Ps_2 (Initial training with R_1)



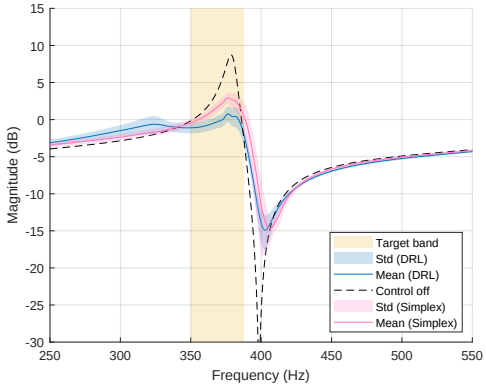
(c) Restart training with R_4



(d) From Pa_2 to Ps_2 (Restart training with R_4)



(e) Continue training with R_4



(f) From Pa_2 to Ps_2 (Continue training with R_4)

Figure 5.22: Reward variation with episodes (a-c-e) and Bode plots (b-d-f) for the 20 best tuned controllers with DRL and simplex methods applied to the model for different strategies and reward definitions.

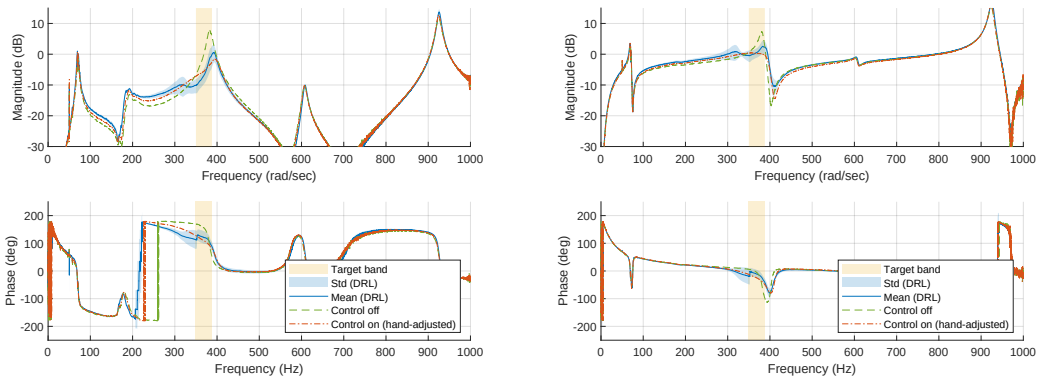
.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

5.3.3 Control performances on the test bench

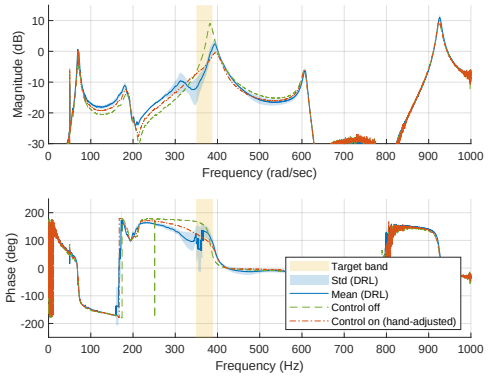
The efficiency of the DRL tuned controllers is evaluated on the experimental test bench. A quantitative analysis is presented which evaluates the damping effect on the aiming mode.

Some tuned controllers have been observed to exhibit experimental instability, with only 10 out of 20 measurements yielding valid results. The reward definition is adapted to the experimental measurement by considering transfer functions from Pa_2 to Ps_1, Ps_2 and Ps_3 as shown in Figure 5.23. Some tested controllers are listed in Table 5.6.



(a) From Pa_2 to Ps_1

(b) From Pa_2 to Ps_2



(c) From Pa_2 to Ps_3

Figure 5.23: Bode plots for 10 tuned controllers with DRL measured on the experimental setup.

- Table 5.6 - Tuned Values: The parameters tuned by the controllers are different from each other. Some gains are null, resulting in unrealistic values for ξ and ω that have no effect on the control of the structure. The tuned references were chosen to be identical for each loop. The DRL tuned controllers highlight the possibility of using only two controllers over three to achieve the same control objective. This phenomenon is interesting from a design point of view to reduce the energy and utility resources needed to build and control the structure.

- Table 5.6 - Reward: The performance of the DRL-tuned controllers varies around the reference value.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Table 5.6: Control Law values and reward computed from measurements on the experimental setup (Selected examples DRL 1, 3, 5 and 10 out of 10 training sessions)

Tuning Method	Tuned Values									Reward
	K_{c1}	K_{c2}	K_{c3}	ξ_1	ξ_2	ξ_3	ω_1	ω_2	ω_3	R_4
Crtl off	-	-	-	-	-	-	-	-	-	9.54
Reference	-4e5	-4e5	-4e5	0.23	0.23	0.23	2387	2387	2387	14.8
DRL 3	0	-1,75e6	-7,00e5	0,80	0,45	0,19	6273	2869	2142	10.62
DRL 1	-1,07e6	0	-6,09e5	0,32	0,78	0,11	2255	1479	2152	10.71
DRL 5	-1,99e6	0	-5,73e5	0,42	0,41	0,11	2881	3143	2169	21.97
DRL 10	-6,82e5	-9,40e5	0	0,21	0,11	0,91	2265	2198	108	11.18

- Figure 5.23: The damping effect of the reference controller is included into the performance STD of the DRL-tuned controllers.

These results show the efficiency of the DRL proposed method to achieve a target control performance. It also provides complementary information on the control of the structure by highlighting the possibility of using only two control loops over three to achieve the same objective.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

5.4 Conclusion

This chapter proposes an efficient method to tune distributive controller parameters using deep reinforcement learning tools on a smart structure.

The structure design by transducer placement is crucial to have high piezoelectric coupling factor and realize control on the aiming mode at 380 Hz. To perform multiple calculations, a transfer function model is estimated from experimental measurements on the test bench. Online training directly on the test bench is time consuming and risks damaging the setup due to piezoelectric instability when the closed-loop diverges. The model has limitations when including feedback control laws and near instabilities. The differences between the model and the experimental setup are significant but improving the model is beyond the scope of this study. Parameters of distributive PPF controllers are adjusted to damp vibrations of the target bending mode. A reference controller is tuned manually to set a realistic control goal.

The Simplex and DRL optimization methods are employed in a numerical environment to tune the controllers, with the objective defined by the reward function. The chosen reward definitions considers H_2 norm minimization of the three colocated transfer loops within a frequency range. It can be criticized because one controller behavior can compensate others.

As demonstrated in Part 5.2.2.3 with random initialization, DRL exhibits superior reliability compared to simplex in achieving a stable solution. Despite requiring more computational resources, it interacts less with the environment, thereby reducing the time spent on interactions with the environment. The use of DRL as a tuning tool remains an effective approach, as it is capable of dampening the target structure modes with the same efficiency as the reference controller. As seen in Chapter 3, to obtain better damping performances and outperform the reference control, DRL agent can be trained with more episodes.

The principal benefit of DRL methods is their capacity to adapt to environmental fluctuations, as evidenced by the objective modification through reward definition. The utilization of a previously trained agent facilitates the achievement of the objective reward value in a more time-efficient manner. It would be beneficial to identify a metric that can quantify the influence of environmental alterations on the rate of adaptation.

To go further, the use of an adaptive model with, for example, a numerical twin could be useful to address the control instabilities not considered in the uncontrolled model. Also, instabilities resulting from uncontrolled modes could be reduced by using other tuned PPF controller combinations.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Chapter 6

Tuning distributed shunt control of a smart structure: numerical study

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

This chapter studies a distributed structure to control with shunted piezoelectric transducers. The free end beam has 20 pairs of collocated transducers distributed along its length (Figure 6.1).

The work presented in this chapter was carried out with the help of Masters students from Ecole Centrale Lyon (Mathis Hamani, Kinane Obeidine and Youssef Ghaddari) and Engineers (Abdelhakim Ezzerouki, Kevin Billon and Corentin Jorajuria). The work of former doctoral students is employed as preliminary work in this research project. A FEM model has already been built with experimental adjustments. The smart beam was studied in the following research projects:

- Fabien Maugan et al. [90] modeled the beam to create a wave trap at the center of the beam. Negative capacitance shunts were used to locally modify the Young's modulus across the beam so that all the energy is focused at a specific location where a harvesting device is used.

- Peng WANG [149]-[150] developed a methodology to obtain a centralized simple MIMO controller. Using FEM, reduction and identification, an accurate low order model of the smart structure is obtained. A simplified H_∞ controller is used to reduce the oscillations in the central zone of the beam with stability guarantee. However, the use of a centralized controller requires a high computational load on the central processor due to its complexity.

- Eva Markiewicz [87] developed a distributive control strategy that considers structural subsystems and uses the LMI method to optimize all the controllers. This method requires less computational resources as fewer variables need to be processed compared to the centralized method. However, a strong hypothesis is used to develop the method, which considers the global structure as stable if all subsystems are stable. This statement seems unrealistic in an experimental context, because each closed-loop introduces small changes into the system and affects the stability of the neighboring loops.

The aim of this chapter is to use the deep reinforcement learning as a centralized algorithm to implement a decentralized strategy with global performance and stability of the control. With 20 control laws to tune, deep reinforcement learning strategies are compared in terms of control performance and computational requirements with genetic algorithms as an optimization method.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

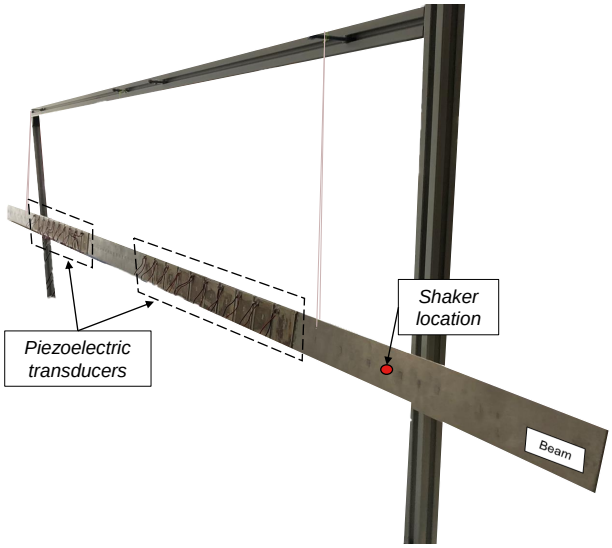


Figure 6.1: Experimental setup: beam with piezoelectric transducers.

6.1 Definition of a smart multi-input multi-output system

The structure under investigation is an aluminium beam containing 40 piezoelectric transducers mounted in pairs. A finite element model has been adapted to the experimental test bench in previous studies. A modal analysis is performed on the structure to define the control objectives and feasibility. A reduced order model is built using modal projection. This reduces the computation time.

6.1.1 Introduction to the study structure

6.1.1.1 Test bench setup

The beam is suspended by wires as shown in the Figure 6.1 to perform almost free end conditions. The beam is divided into two section types along the length axis: the classical material and the metamaterial where the transducers are located. Disturbances are introduced into the beam by a shaker (not shown on the picture).

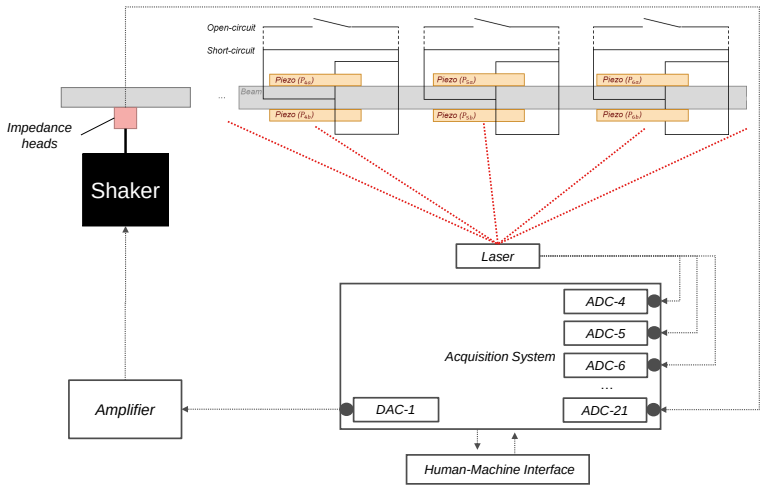


Figure 6.2: Experimental setup: data acquisition systems

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

6.1.1.2 Test bench measurements

Preliminary measurements are made on the beam to determine its modal density. A noise with a sampling frequency of $Fs = 2000$ Hz is applied to the shaker F_0. Displacement measurements are carried out using a Laser Scanning Vibrometer at the center of each transducer location $P_{i=1;...;20}$. Figure 6.3 shows the transfer function between the shaker force and the displacement at transducer location P_1, as an example, with all transducers set to short-circuit or open-circuit conditions.

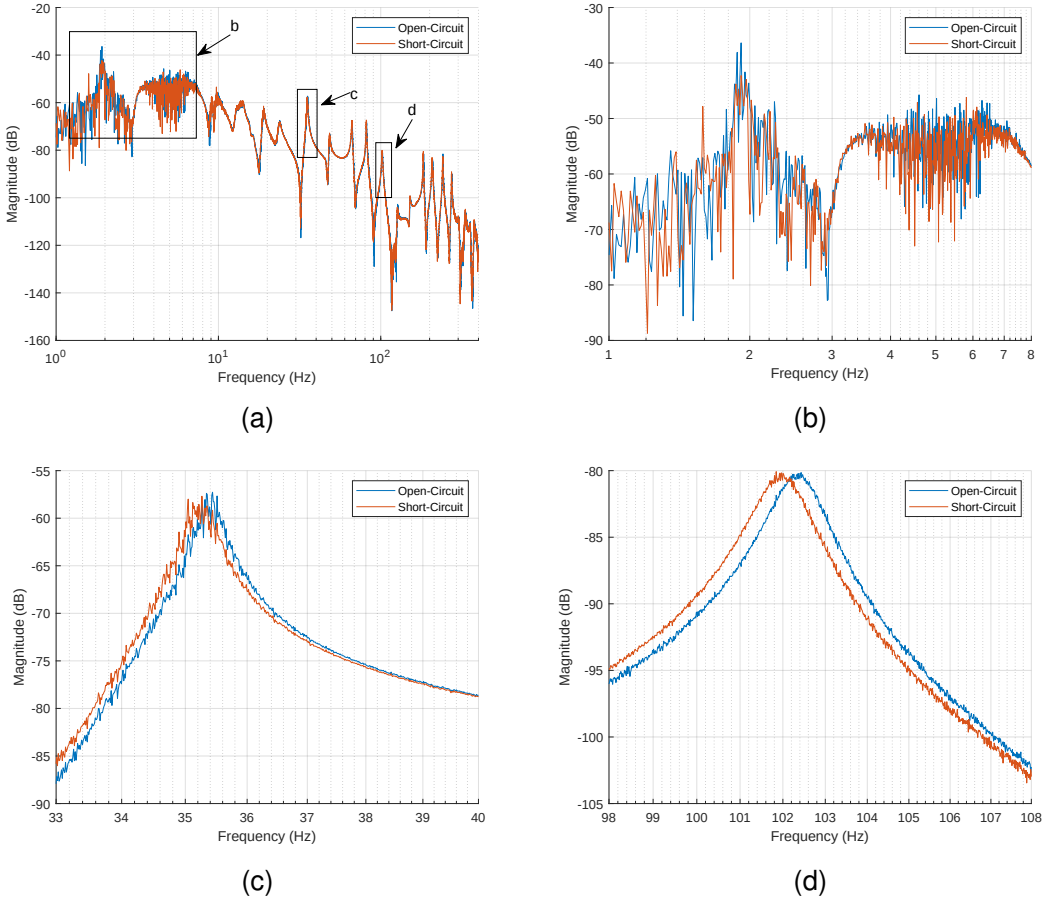


Figure 6.3: Bode from the force F_0 at the shaker location to the displacement measured at transducer P_1 location.

Experimental modal parameters are extracted from the measurement using Baptiste Chomette MIMO Modal Parameters Identification in Frequency Domain [18] and reported in the Table 6.1. The coupling coefficients are reaching almost 10% on this structure, as in the previous chapters where the coupling coefficient was close to 15%. The effective modal damping is also heterogeneous with frequency and varies from 0.35% to 0.96%. The structure energy response is mostly focused at low frequencies above 10 Hz (Fig.6.3a and 6.3b). Modal identification could not be performed at low frequencies because the signal coherence is too low due to the limitations of the shaker and the wire suspensions. However, the control strategy should focus on damping the most energetic part of the vibration spectrum at low frequency. Ideally, the control should be able to cover a wide frequency range to attenuate higher frequencies.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Table 6.1: Modal parameters estimated from measurements from Figure 6.3 transfer function

Open-Circuit		Short-Circuit		Coupling coefficient
Frequency (Hz)	Damping ξ	Frequency (Hz)	Damping ξ	K_{eff} (%)
35,32	0,59	35,22	0,56	7,47
48,35	0,76	48,18	0,68	8,31
66,70	0,57	66,37	0,55	9,91
81,84	0,44	81,54	0,45	8,53
102,36	0,66	101,98	0,68	8,68
126,82	0,74	126,41	0,69	8,12
151,58	0,59	151,02	0,59	8,63
183,59	0,35	182,77	0,34	9,41
208,38	0,53	207,36	0,55	9,87
241,80	0,36	240,87	0,36	8,77
274,00	0,35	272,89	0,35	8,98
303,74	0,89	303,17	0,96	6,15
357,01	0,61	356,60	0,63	4,80

6.1.2 Model definition of the smart beam

A finite element model was constructed and to fit in frequency with the experimental results, but without considering damping adjustments as they do not affect the optimization strategy. The model is reduced using a state space representation with modal projection between 0 and 590 Hz.

6.1.2.1 Finite element model definition

The beam dimensions and material properties are given in Table 6.3 with the variable names defined in Figure 6.4. Experimental capacities refereed in Table 6.2 are implemented in the model. As they are arranged in parallel, the equivalent capacity is calculated as follows:

$$Cp^i_{eq} = Cp^i_a + Cp^i_b \tag{6.1}$$

Where Cp^i_{eq} is the equivalent capacitance of transducers $i = 1; ...; 20$, Cp^i_a is the capacitance of the transducer i at surface (a) and Cp^i_b at surface (b).

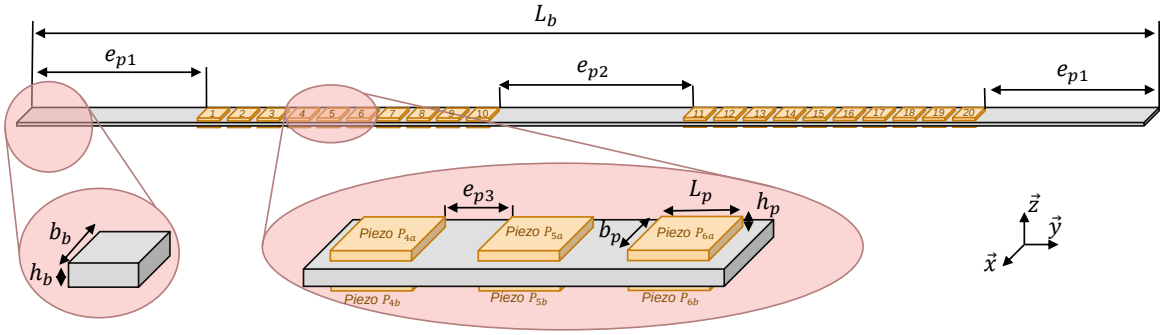


Figure 6.4: Beam dimensions diagram

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Table 6.2: Transcuders capacitance: measurements

Transducers	Face a [nF]	Face b [nF]	Equivalent [nF]
1	43.6	48.4	92.0
2	48.4	47.8	96.2
3	48.5	48.1	96.6
4	47.5	48.0	95.5
5	47.3	48.3	95.6
6	47.4	48.4	95.8
7	48.8	49.1	97.9
8	47.7	48.7	97.4
9	48.9	49.2	98.1
10	48.5	48.3	96.8
11	47.6	48.5	96.1
12	47.8	48.5	96.3
13	47.7	48.4	96.1
14	47.8	48.1	95.9
15	48.8	48.2	97.0
16	48.5	49.1	97.6
17	48.3	48.3	96.6
18	48.7	48.9	97.6
19	48.4	49.0	97.4
20	48.1	49.5	97.6

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Table 6.3: Definition of the constants

	Variable	Beam (Aluminum)		Piezoelectric transducer (PZT-PIC255)		Unit
		Name	Value	Name	Value	
Dimensions	Length	L_b	2.5	L_p	0.05	$[m]$
	Width	b_b	0.053	b_p	0.05	$[m]$
	Thickness	h_b	0.003	h_p	0.001	$[m]$
	Location	-		e_{p1}	0.45	$[m]$
		-		e_{p2}	0.41	$[m]$
		-		e_{p3}	0.01	$[m]$
Material properties	Young modulus	E_b	70e9	c_{11}^S	1.20e11	$[Pa]$
				c_{12}^S	7.52e10	$[Pa]$
				c_{13}^S	7.50e10	$[Pa]$
				c_{33}^S	1.11e11	$[Pa]$
				$c_{44}^S = c_{55}^S$	2.11e10	$[Pa]$
				c_{66}^S	2.58e10	$[Pa]$
	Density	ρ_b	2700	ρ_p	7700	$[kg.m^{-3}]$
	Coupling constant	-		e_{31}	-2.8	$[C.m^{-2}]$
		-		e_{33}	14.7	$[C.m^{-2}]$
		-		e_{15}	9.86	$[C.m^{-2}]$
	Relative permittivity	-		ϵ_{11}^S	828	$[-]$
		-		ϵ_{33}^S	700	$[-]$

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

6.1.2.2 Model building

A state space model is built from the previously defined Finite element Model considering N_b nodes with the same equations as in Chapter 2 Part 1.1.3:

$$\begin{bmatrix} M & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \ddot{w} \\ \ddot{V} \end{bmatrix} + \begin{bmatrix} K & K_{me} \\ -K_{me}^T & C_p \end{bmatrix} \begin{bmatrix} w \\ V \end{bmatrix} = \begin{bmatrix} F \\ q \end{bmatrix} \tag{6.2}$$

Matrix are noted $M \in \mathbb{R}^{2N_b \times 2N_b}$ for the mass, $K \in \mathbb{R}^{2N_b \times 2N_b}$ for the stiffness, K_{me} for the piezoelectric component, C_p for the coupling coefficient $V \in \mathbb{R}^{20}$ for the voltage and $q \in \mathbb{R}^{20}$ for the electric charges on each piezoelectric transducer, $F \in \mathbb{R}^{2N_b \times 1}$ for the external forces applied on the system, and $w \in \mathbb{R}^{2N_b}$ for the displacements w_j and θ_j of the system, such as:

$$\{w\} = \begin{bmatrix} w_1 \; \theta_{\,1} \; \ldots \; w_{\mathrm{N_b}} \; \theta_{\,\mathrm{N_b}} \end{bmatrix}^T \tag{6.3}$$

Using modal projection on the 17th first bending modes defined between 0 and 590 Hz, the modal vector η is defined, such as:

$$\{w\} = \Phi\eta \tag{6.4}$$

Where Φ is the coordinates matrix from the physical space to the modal space solution to the eigenvalue problem $(-\Omega_i^2 M + K)\Phi_i = 0$.

$$\begin{bmatrix} \Phi^T M \Phi & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \ddot{\eta} \\ \ddot{V} \end{bmatrix} + \begin{bmatrix} \Phi^T K \Phi & \Phi^T K_{me} \\ -K_{me}^T \Phi & C_p^* \end{bmatrix} \begin{bmatrix} \eta \\ V \end{bmatrix} = \begin{bmatrix} \Phi^T F \\ q \end{bmatrix} \tag{6.5}$$

Modal truncation does not allow the quasi-static behavior of the electrostatic equilibrium to be recovered. A static correction is made to the position of the poles and zeros of the electrical impedance, i.e. that of the initial system, which guarantees a good assessment of the modes at zero currents and voltages:

$$C_p^* = C_{stat} - Kme^T \phi diag(M\Omega^2)^T \phi^T Kme \tag{6.6}$$

Divided the first line by the modal mass matrix :

$$\begin{bmatrix} Id & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \ddot{\eta} \\ \ddot{V} \end{bmatrix} + \begin{bmatrix} \Omega^2 & \Phi^T K'_{me} \\ -K_{me}^T \Phi & C_p^* \end{bmatrix} \begin{bmatrix} \eta \\ V \end{bmatrix} = \begin{bmatrix} \Phi^T F' \\ q \end{bmatrix} \tag{6.7}$$

A state space representation of the system is built, such as:

$$H = \begin{cases} \dot{X} = \mathrm{AX} + \mathrm{BU} \\ Y = \mathrm{CX} + \mathrm{DU} \end{cases} \tag{6.8}$$

Where $X = \begin{bmatrix} \eta \\ \dot{\eta} \end{bmatrix}$ is the state vector, $U = \begin{bmatrix} F_t \\ V \end{bmatrix} \in \mathbb{R}^{40}$ is the input vector with $F_t \in \mathbb{R}^{20}$ the force at each transducer location, $Y = \begin{bmatrix} q \\ y \end{bmatrix} \in \mathbb{R}^{60}$ is the output vector with $y \in \mathbb{R}^{40}$ the displacement w_j and θ_j at each transducer location and matrices A, B, C and D are defined, such as:

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

$$A = \begin{bmatrix} 0 & \mathrm{Id} \\ -\Omega^2 & -2\xi\Omega \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 \\ \Lambda^T & -\Phi^T K_{me} \end{bmatrix}$$
$$C = \begin{bmatrix} -K_{me}^T \Phi & 0 \\ \Lambda^T & 0 \\ 0 & \Lambda^T \end{bmatrix} \quad D = \begin{bmatrix} 0 & C_p \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Where Λ^T it the transformation matrix from the modal basis to the transducers coordinate and ξ is the damping coefficient. Note that the damping coefficient does not match the experimental test bench in the following part. Also, the model with a non-zero D matrix has infinite H_2 energy.

6.1.2.3 Modal coupling

The modal response of the beam is studied using the models. The modal shapes are classified in Appendix D between 0 and 590 Hz as follows:

- Rigid modes: 6
- Bending modes in y-axis: 5
- Torsional modes in x-axis: 6
- Z-axis bending mode: 21

The modes identified during the measurements are only bending in the z-axis as shown in Figure 6.5.

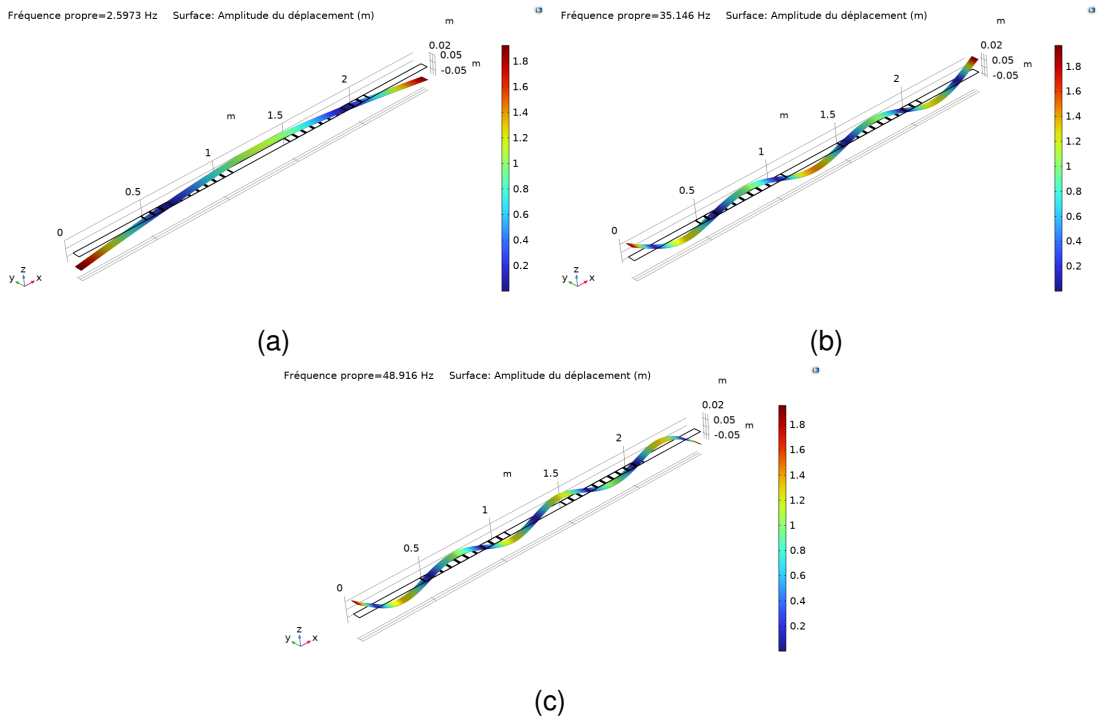


Figure 6.5: Mode shape visualization: first bending mode (a), fifth bending mode (b) ninth bending mode (c).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

The modal parameters extracted from the model are shown in Table **??** and compared with the experimental ones. As expected, the model fits the experiment according to the frequencies, with absolute errors ranging from 0.09 to 4.5 % depending on the mode. However, the coupling coefficient is twice overestimated, reaching 14.67 % for the mode at 35.35 Hz, where it is estimated from measurements to be 7.74 %. This difference may be explained by the use of a polymer to bond the transducer to the beam. This layer is not modeled and its thickness can vary due to human manipulation.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Table 6.4: Modal parameters estimated from the model and compared with measurements from Table **??**

Open-Circuit Frequency (Hz)	Short-Circuit Frequency (Hz)	Coupling coefficient K_{eff} (%)	Open-Circuit Absolute error (%)	Short-Circuit Absolute error (%)	Coupling coefficient Absolute error (%)
2,85	2,82	13,70	-	-	-
7,94	7,80	18,86	-	-	-
14,47	14,32	14,33	-	-	-
23,05	22,84	13,72	-	-	-
35,35	34,98	14,67	0,09	0,69	49,13
48,78	48,28	14,44	0,89	0,20	42,49
66,48	65,62	16,22	0,33	1,14	38,89
83,58	82,70	14,58	2,07	1,40	41,50
104,40	103,27	14,86	1,95	1,25	41,60
127,96	126,63	14,53	0,88	0,17	44,17
153,51	151,86	14,78	1,25	0,55	41,58
184,43	182,18	15,76	0,45	0,32	40,32
212,60	210,28	14,89	1,98	1,38	33,74
246,18	243,35	15,28	1,77	1,02	42,65
279,67	276,73	14,61	2,02	1,38	38,55
318,10	314,53	15,10	4,51	3,61	59,30
361,60	357,28	15,59	1,26	0,19	69,19

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

6.1.3 Implementing control

In this structure, the vibration is controlled by an electrical shunt on the piezoelectric transducers. Each shunt Z_j is independent of the others as shown in Figure 6.6.

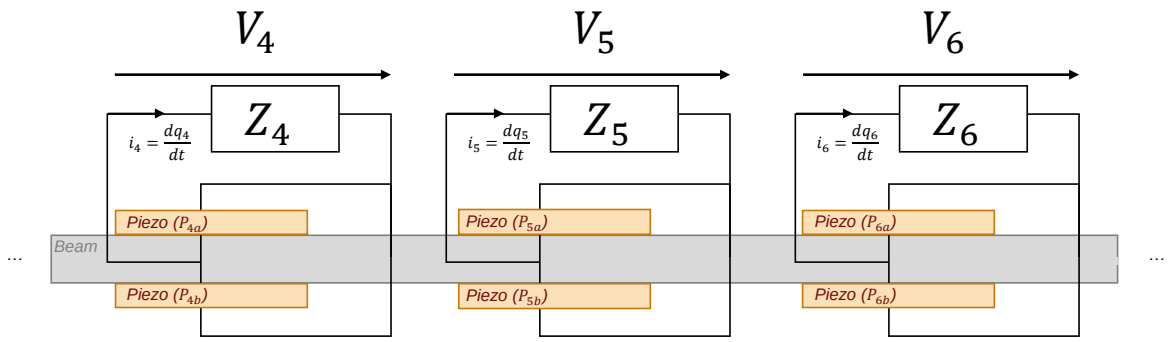


Figure 6.6: Shunt Scheme

The tension V_j at each transducer location is written such as:

$$V_i = -Z_j i_j = -Z_j \frac{dq_j}{dt} = -Z_i s q_i \tag{6.9}$$

Where s is the Laplace variable, i_j is the current and q_j is the electric charge at the transducer j.

Considering a MIMO system, the control can be represented as a block diagram with a feedback loop as shown in Figure 6.12. The controller is written as:

$$G(s) = Zsq = \begin{bmatrix} Z_1 s & 0 & \dots & 0 \\ 0 & Z_2 s & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & Z_{20} s \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_{20} \end{bmatrix} \tag{6.10}$$

Note that the current i is implemented as a function of V to put the roll off in the control algorithm.

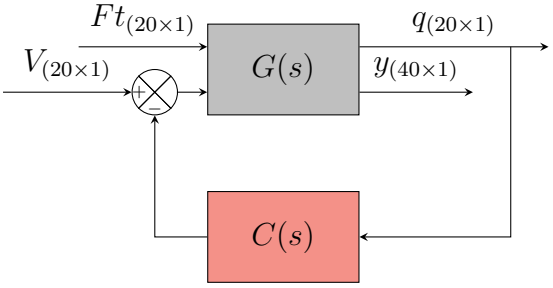


Figure 6.7: Control loop block diagram

According with the literature (Chapter 2, Part 2.1.2.2), it is chosen to focus on two type of shunt:

- Resistive shunt: $Z_j = R_j$. This shunt is optimally effective in a monomodal context.
- Capacitive shunt: $Z_j = R_j + \frac{1}{C_j s}$ where C_j can be negative. This shunt is effective in a multimodal context even if it is tuned to damp vibration of a specific mode due to the dynamics of the beam (predominance of the first bending mode due to free boundary conditions).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

6.2 Control tuning methods

The objective is to find the appropriate shunt values to control the beam vibrations. Different methods are used and their performance is compared to the DRL method in terms of computational resources and control efficiency. As in the previous chapter, this optimisation is performed on a defined environment consisting of the structure, the controller parameters and the control objective represented by the reward.

6.2.1 Environment definition

The environment is based on the state space model $G(s)$ previously defined in this chapter.

6.2.1.1 Input disturbance

The input disturbance is introduced to the system through a force at the location of the first transducer P_{a1}. It is defined by a noise signal v_d with a sampling frequency $F_s = 2500$ Hz, such as:

$$F_t = \begin{bmatrix} v_d = Rnd(t) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{6.11}$$

Where t is the time. The simulation is run for 10 seconds. In this configuration the perturbation is chosen to be collocated with a transducer used for control, but the methods are the same with a force introduced on the beam at a different location. The only difference is in the modal response of the structure according to the node positions of the modes. Here, for modes at 49, 158 and 190 Hz (6th, 11th and 12th bending mode), the input perturbation is close to a node, as can be seen from the modal shapes in the Appendix D.

6.2.1.2 Observations

The observations are defined as the root mean square of the displacement, such as:

$$S = \begin{bmatrix} S_1 \\ S_2 \\ ... \\ S_{20} \end{bmatrix}_{v_d} = \begin{bmatrix} rms(w_1) \\ rms(w_2) \\ ... \\ rms(w_{20}) \end{bmatrix} \tag{6.12}$$

The displacements are chosen as observations to mimic a non-invasive measurement method that can be performed in a test bench configuration with a laser. To avoid large values as NN input, the values are normalized between -1 and 1 over S.

6.2.1.3 Action

Two action matrices can be defined, taking into account the different control by shunt:

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

- Resitive shunt: $A_R \in \mathbb{R}^{20}$

$$A_R = log10 \begin{bmatrix} R_1 \\ R_2 \\ ... \\ R_{20} \end{bmatrix} \tag{6.13}$$

- Capacitive shunt: $A_{RC} \in \mathbb{R}^{20\times 2}$

$$A_{RC} = log_{10} \begin{bmatrix} R_1 & -C_1 \\ R_2 & -C_2 \\ ... & ... \\ R_{20} & -C_{20} \end{bmatrix} \tag{6.14}$$

To avoid large numbers as NN output, the actions are defined by the logarithm of the electrical component values. The limit of the resitive values are chosen as $log_{10}(R_{shunt}) \in \{1; ...; 8\}$. For the negative capacitance, the limits are chosen according with the transducer capacitance as $log_{10}(-C_{shunt}) \in \{-10; ...; -6\}$. This range of negative capacitance values includes an unstable domain (as seen in Chapter 2, Part 1.2.2) for values close to the piezoelectric capacitance of the transducer.

6.2.1.4 Reward

The purpose of the reward definition is to quantify the performance of the shunt applied while ensuring the stability of the system. A conditional definition is chosen such as:

- Stable system :

$$Reward = 200 \times (log_{10} \frac{1}{20} \sum_{j=1}^{20} \frac{rms(wj)^{off}}{rms(wj)^{on}}) \tag{6.15}$$

  This definition is designed such that the short-circuit case results in a null reward. A multiplier coefficient of 200 is chosen to increase reward sensitivity.

- Unstable system or large negative rewards :

$$Reward = -10 \tag{6.16}$$

  Some shunt configurations can lead to large negative rewards or unstable cases. Setting a low limit avoids training the NN with useless precise data.

Figure 6.8 shows the reward variation in the case of a resistive shunt only. To provide a 2-dimensional representation, the same value is defined for all shunt resistances. Here only the stable case with resistive value variations occurs. The resulting function is continuous and has a global maximum.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

The reward is an accurate representation of the expected behaviour:

- (1) The reward increases with resistance
- (2) The optimal resistance is reached
- (3) The reward decreases while the resistance is large enough to be assimilated to an open circuit case and even amplify a little the vibrations.

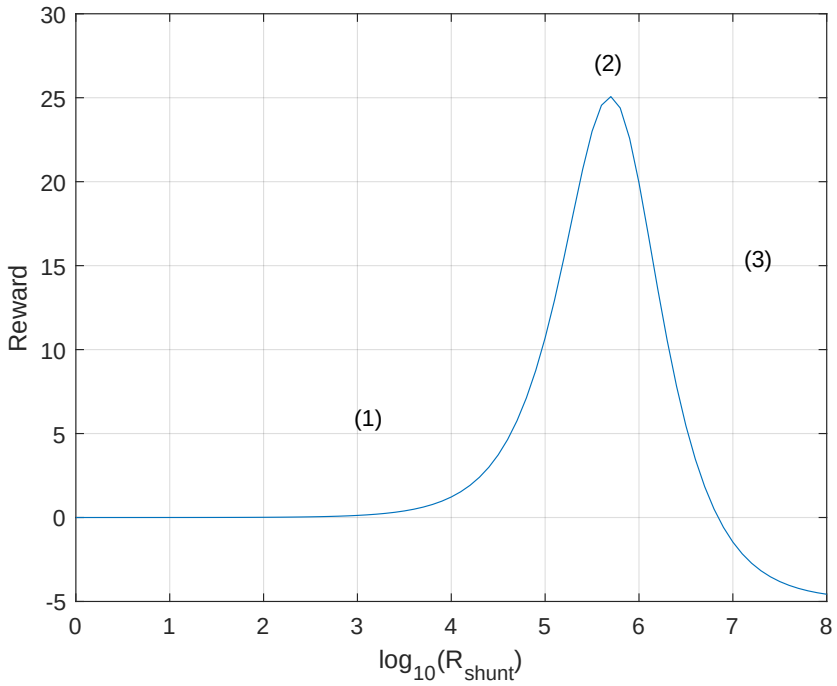


Figure 6.8: Reward variations with Resistive Shunt

Figure 6.9 shows the reward variation in the case of a negative capacitance shunt. The same shunt parameter value is used for all transducers. As seen in the previous behaviours, the reward function still has a global maximum value according to the resistance variations for a given capacitance. The reward variation are mostly influenced by the negative capacitance values :

- (1) Unstable area: $C_{neg} < -10^{-7}$: The negative reward indicates a vibration amplification leading to an unstable system when it reaches -10.
- (2) Optimal area: $-10^{-7} < C_{neg} < -10^{-6.8}$: The reward increases with the negative capacity value, up to a global maximum.
- (3) Sub-optimal area: $-10^{6.8} < C_{neg}$: The reward decreases as the negative capacity increases until the capacity value is negligible with respect to the resistance value.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

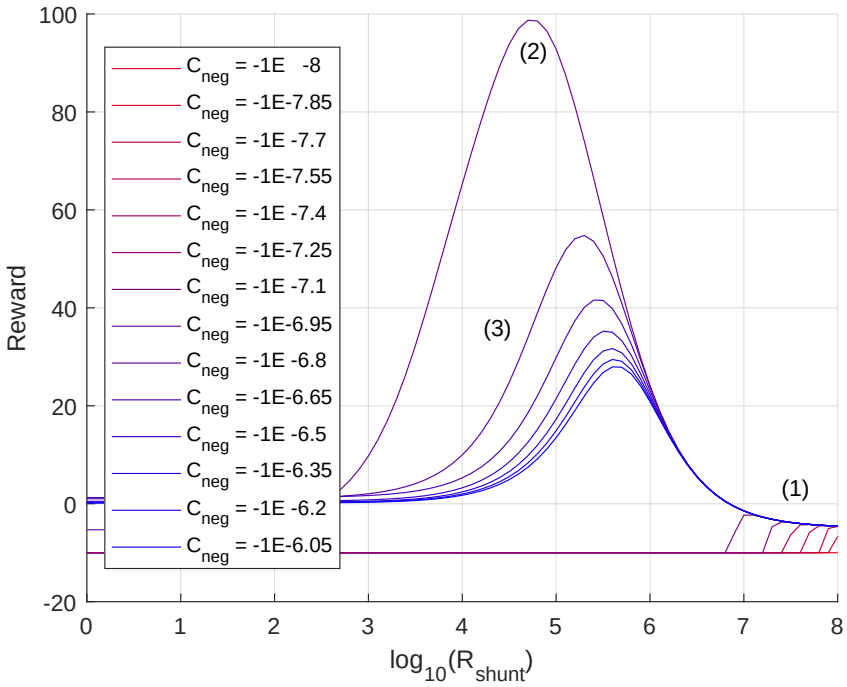


Figure 6.9: Reward variations with $R - C_{neg}$ Shunt

6.2.2 Methods comparison

The objective is to determine the relevance of using deep reinforcement learning tools in the context of a large number of parameters to be tuned: 20 for the resistive or 40 for the capacitive shunt implemented on the structure. It is proposed to compare this new approach with an analytical tuning and a genetic algorithm.

6.2.2.1 Analytic

As a first approximation, analytical tuning from the literature can be used. To apply it to distributive transducers case, the following hypothesis is made:

Hypothesis 5 *The 20 transducers are considered as one with the same equivalent capacitance, calculated as the average of all individual capacitances, such as:*

$$Cp_{eq} = \frac{1}{20} \sum_{j=1}^{j=20} C_{pj} \tag{6.17}$$

All of the patches are connected to independent, but identical, circuits. This hypothesis works in the context where all shunts are set as identical on each piezoelectric transducer. There, the short f_{sc} and open circuit f_{oc} frequencies are evaluated considering all transducer contributions, leading to the evaluation of a global coupling coefficient K_{eff}. In order to tune the optimum shunt to minimize vibration, the following equations can be found in the literature (Chapter 2 part 1.2.2):

- Resistive shunt only: $R_{opt} = \frac{1}{\gamma C p_{eq}(1-K_{eff})}$

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Where $\gamma = f_{oc}\sqrt{\frac{f_{oc}}{f_{sc}}}$ and $K_{eff} = \frac{f_{oc}^2 - f_{sc}^2}{f_{sc}^2}$ is the effective modal coupling coefficient. This optimum resistance value is modal dependent and decreases as the mode frequencies increase, as shown in Figure 6.11.

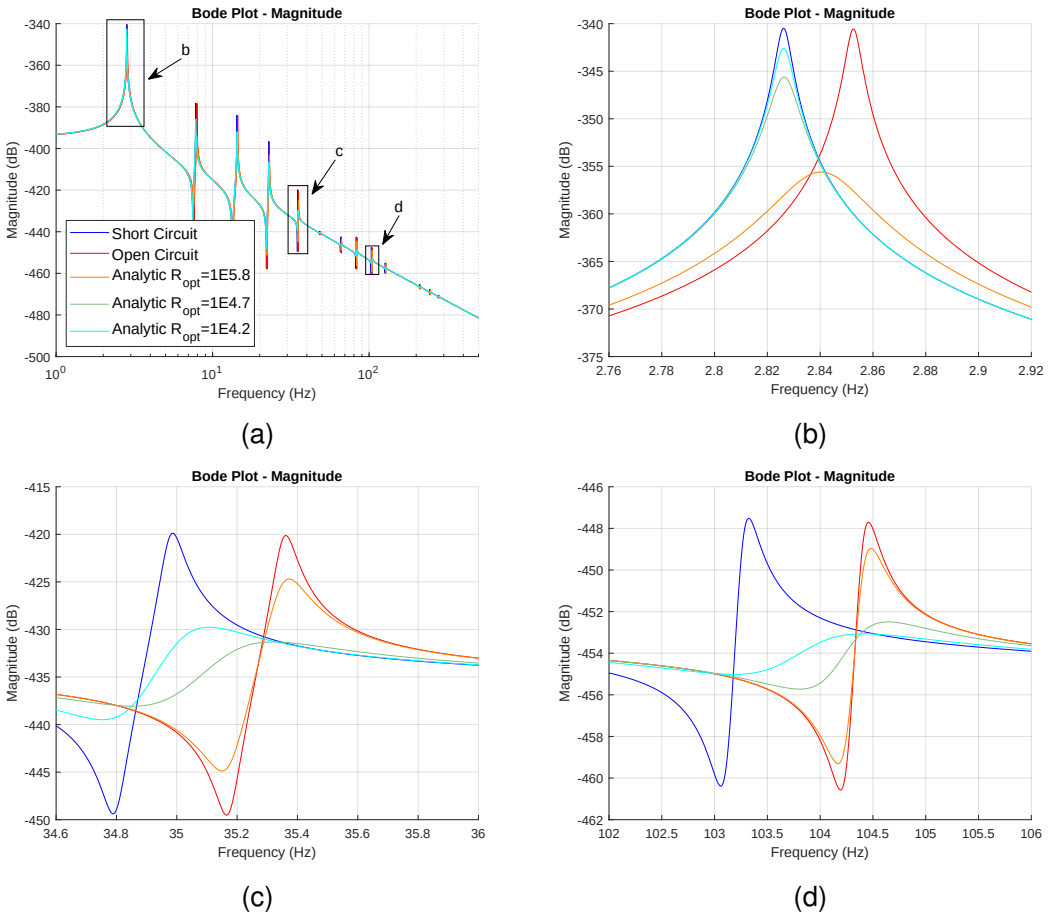


Figure 6.10: Bode R same Analytic tuning

The system response shown focuses on bending modes 1 (Fig.6.10b), 5 (Fig.6.10c) and 9 (Fig.6.10d). These achieve maximum attenuation due to their respective optimum resistive shunts.

- Capacitive shunt: $C_{opt} = -\alpha C p_{eq}$ and $R_{opt} = \frac{\sqrt{2K_{eff}^{Cneg}}}{(1+K_{eff}^{Cneg})Cp_{eq}^{Cneg} f_{sc}^{Cneg}}$

Where $\alpha = 1.1$ is a coefficient set to keep the system stable, f_{sc}^{Cneg} is the short circuit frequency induced by the negative capacitance, K_{eff}^{Cneg} is the new effective coefficient, $Cp_{eq}^{Cneg} = \frac{1}{\frac{1}{Cpeq} + \frac{1}{C_{opt}}}$ is the equivalent capacitance of the transducer (Appendix E provides a demo). The coefficient $\alpha < 1$ guarantees the negative capacitance to induce a softening effect on the structure [81]-[98]. To increase damping performance, α should be as close to 1 as possible, while maintaining stability. Figure 6.11 shows the effect of negative capacitance on the system.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

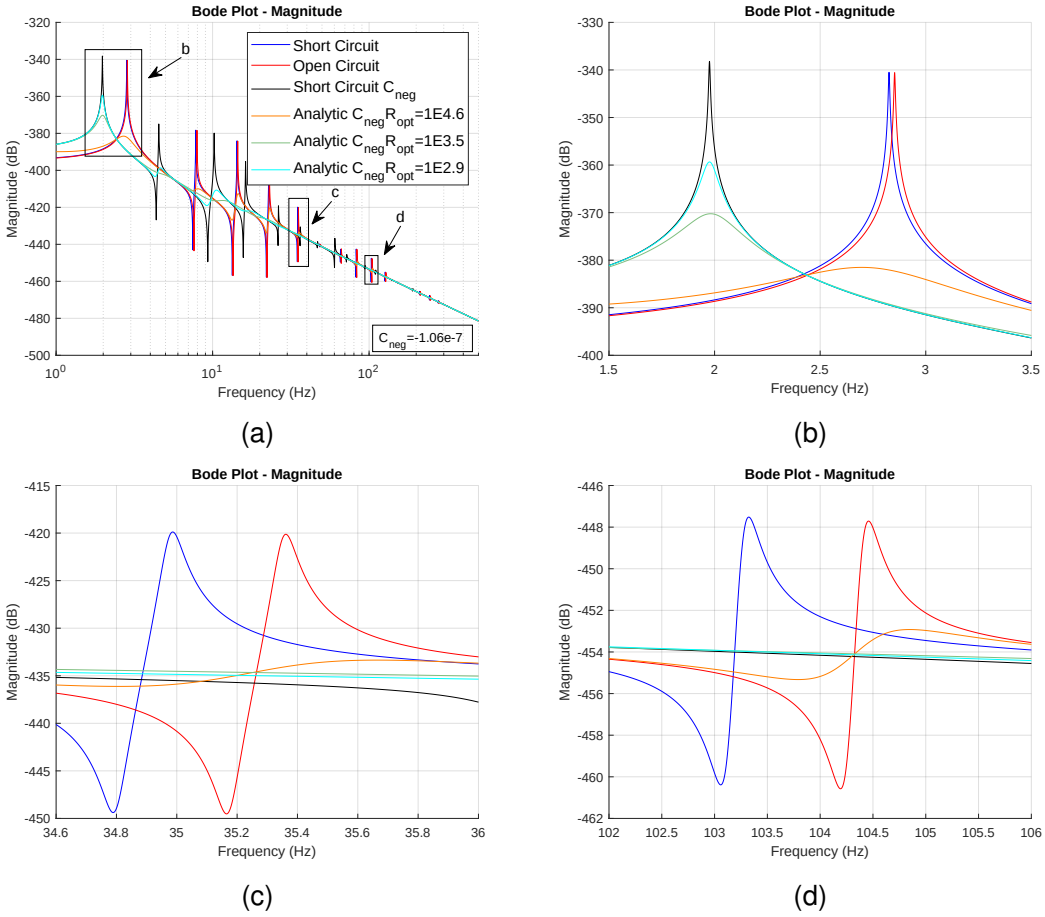


Figure 6.11: Bode RC_{neg} same Analytic tuning

The use of a negative capacitance acts as if the coupling coefficient is artificially increased with a short circuit frequency reduced to lower frequencies [81]. The negative capacitance shunt introduces attenuation throughout all the frequency range of interest. The resistance value adjusts the damping and avoids instability caused by the negative capacitance alone (black line in Fig.6.11).

The tunings obtained with the analytical formulations are used as reference in the following sections to damp the first mode:

- Resistive Shunt only: $R_{shunt} = 1E5.8$ [Ω]
- Capacitance Shunt: $C_{shunt} = -1.1Cp_{eq}$ [F] and $R_{shunt} = 1E4.6$ [Ω]

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

6.2.2.2 Genetic algorithm

Inspired by the concept of biological evolution, genetic algorithms can be used to optimize functions [39]. It is appropriate to use them when the objective functions are non-continuous. The genetic algorithm interactions with the environment is shown on Figure 6.12. Since this algorithm works with minimization, the objective function is chosen as the reward definitions with a negative sign from Part 6.2.1.4. The inner workings of the algorithm are detailed in Chapter 2 Part 1.3.4, and here $m \in \mathbb{N}$ is the number of individuals in the population and $n \in \mathbb{N}$ is the number of generations.

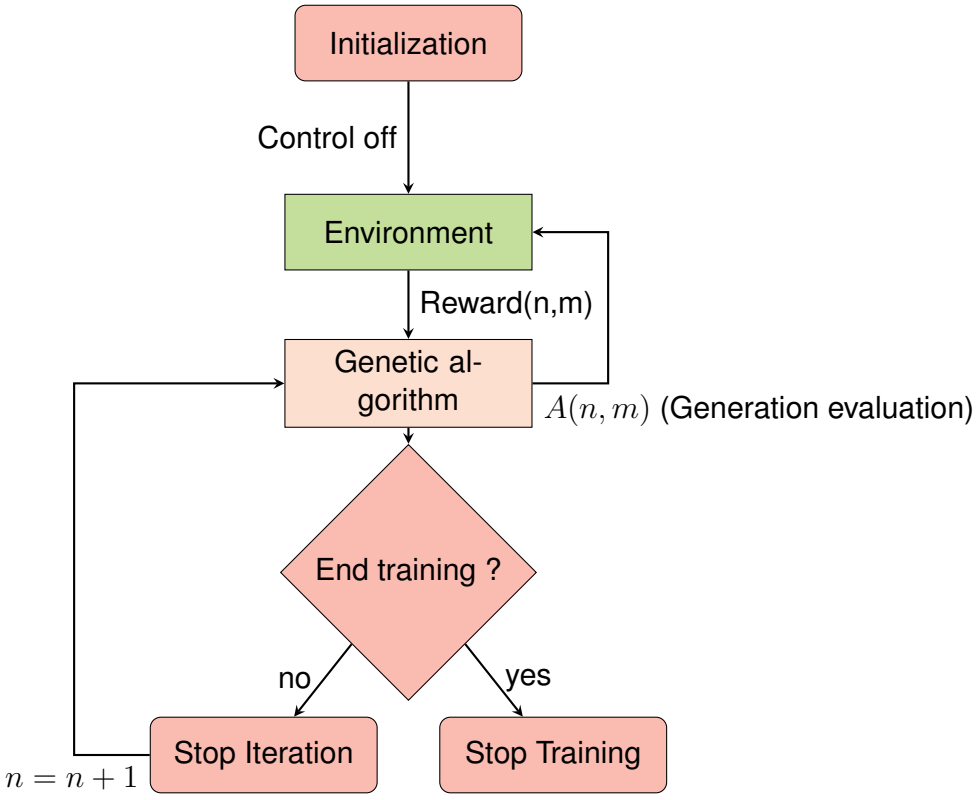


Figure 6.12: Genetic Algorithm diagram

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

6.2.2.3 Deep reinforcement learning

The deep reinforcement learning strategy used is similar to the previous chapters. The TRPO methods is used with the defaults hyper-parameters sets. The agent composed of the actor and the critic neural networks with dimensions defined in Table 6.5. The networks are considered to be sufficiently complex to represent the required pattern between input and output.

Table 6.5: Actor and critic neural network definition

Layer	Name	Number of Neurons Actor	Critic
Input	N_i	20	20
Hidden Layer 1	N_{hu}	25	25
Hidden Layer 2	N_{hu}	25	25
Hidden Layer 3	N_{hu}	25	-
Output	No	1 (R shunt identical case)	1
		20 (R shunt distinct case)	
		2 (RC_{neg} identical shunt case)	
		40 (RC_{neg} distinct shunt case)	

In this chapter, the memory properties of NN is used to perform the training with different steps as shown in Figure 6.13. The objective is to start the training with low values to tune considering all shunts as identical, but independent from each others (DRL0 training). Then the last NN layer is duplicated to fit with the distinct shunt configuration. In this way, the layers keep the memorized pattern. Training can be performed by adjusting each shunt independently (DRL1 training) until the target reward is reached, corresponding to the analytical tuning. The NN can be trained further if the analytical tuning does not correspond to the "optimal" solution (DRL2 training). This method is compared with the one previously presented (DRL3 training) without pre-training. The NNs are set as identical to the initialised agent. A duplication of the last layer is done to perform the training with a different shunt configuration.

6.2.3 Optimisation results on the modelled beam

The optimisation methods are used to tune a resistive (R) or a resistive-capacitive (RC_{neg}) shunt. Two shunt configurations are investigated:

- Identical shunt case: the same shunt value is applied to all transducers but all are independent.
- Distinct shunt case: the shunt value applied to all transducers is independent

The algorithm are all run on the above defined environment (part 6.2.1).

6.2.3.1 Control with Resistive shunt on the model

Deep reinforcement learning algorithms are trained according to Figure 6.13. The training monitor is shown in Figure 6.14. Table 6.6 gives a comparison of the tuning and performance of all the methods used according to the resistive shunt configuration.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

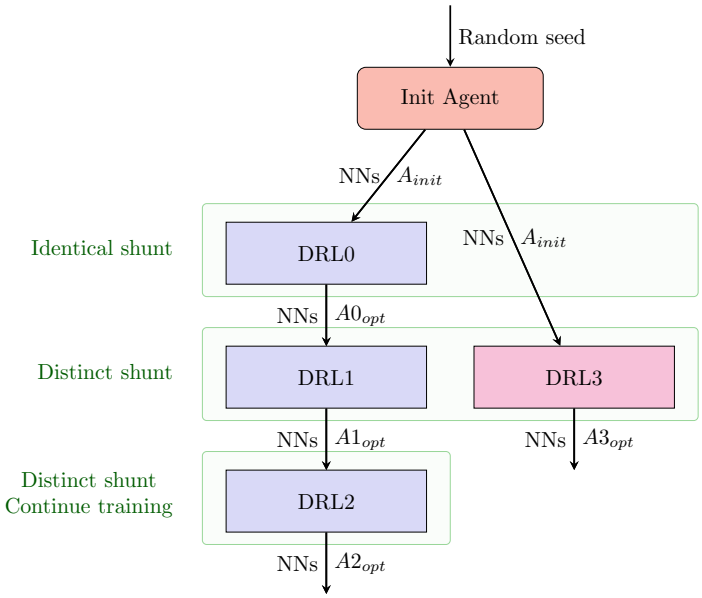
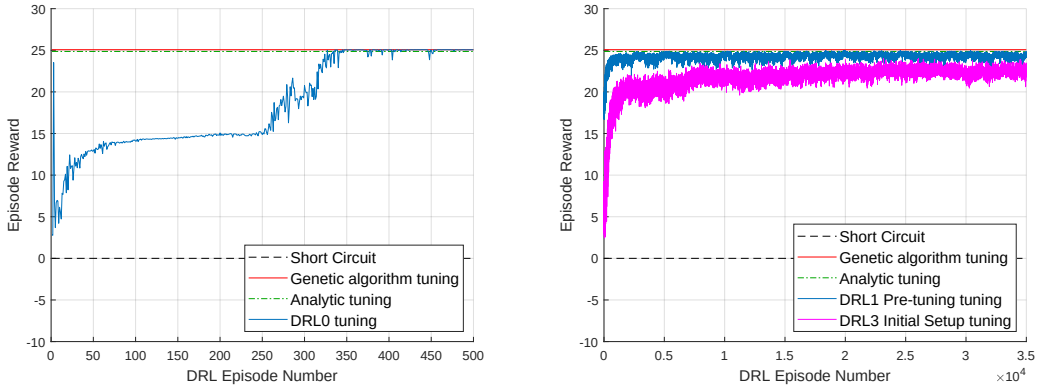


Figure 6.13: DRL Training Process



(a) (b)

Figure 6.14: Training Identical case with DRL0 (a), Distinct case with DRL1 and DRL3 (b): R shunt

For the identical case, DRL0 training tunes a resistance parameter using 500 episodes (Fig. 6.14a), while the genetic algorithm needs 3014 estimations to achieve the same result (Tab. 6.6). Both methods outperform the analytical result (Part 2.2.1) in terms of reward. However, the difference between the three tuning is not noticeable when looking at the system response in Figure 6.15.
For the distinct case, the genetic algorithm needs 35,000 estimates. The DRLs are trained on the same number of episodes. Figure 6.14b shows the training monitoring considering the pre-trained NNs (DRL1) and the initial NNs (DRL3). DRL1 managed to reach the objective reward (calculated from the analytical tuning) in less than 5000 episodes, while DRL3 did not reach it after 35,000 episodes.
The rewards of the distinctive and identical shunt cases are very similar as the structure response from Figure 6.16. The first bending mode of the structure has the greatest influence on its response, which explains why the shunts are tuned to minimize it in both the identical and different cases.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

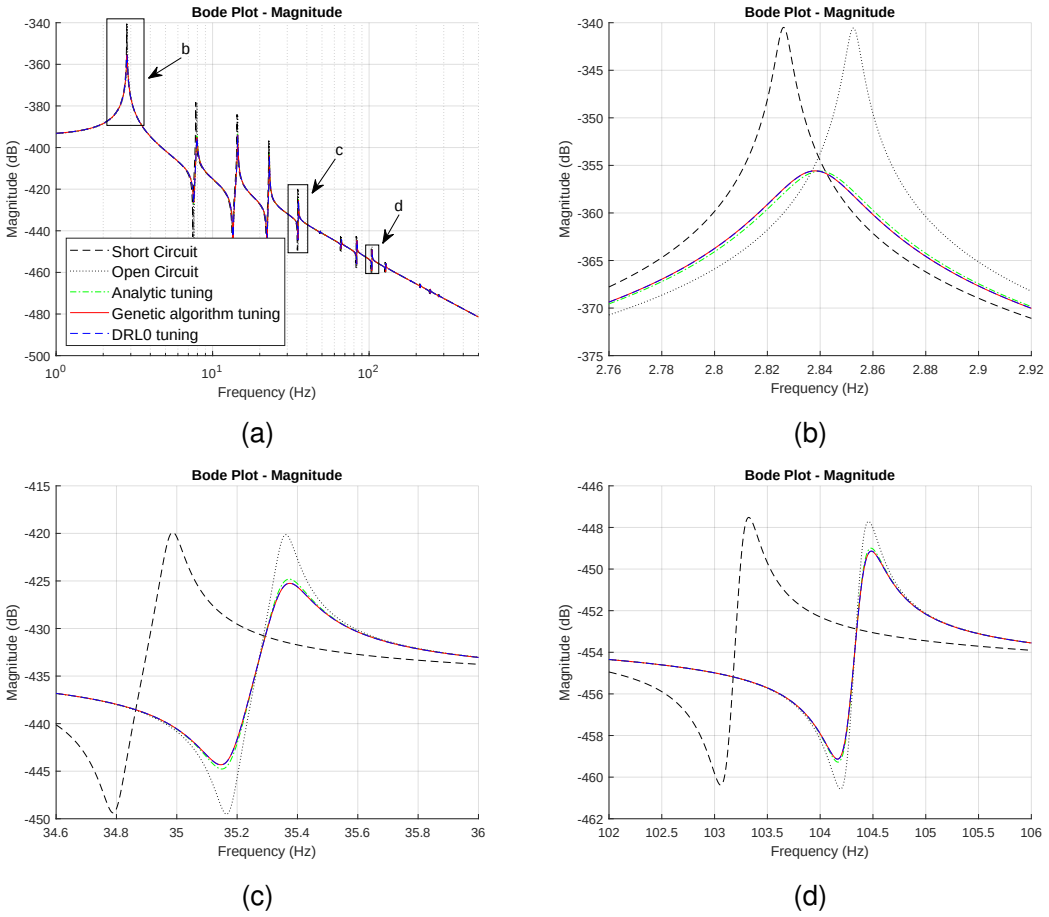


Figure 6.15: Bode from F_{t1} to w_1 (a), with zoom on the first mode (b), the fifth (c) and the ninth (d) bending mode: R shunt - Identical case

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

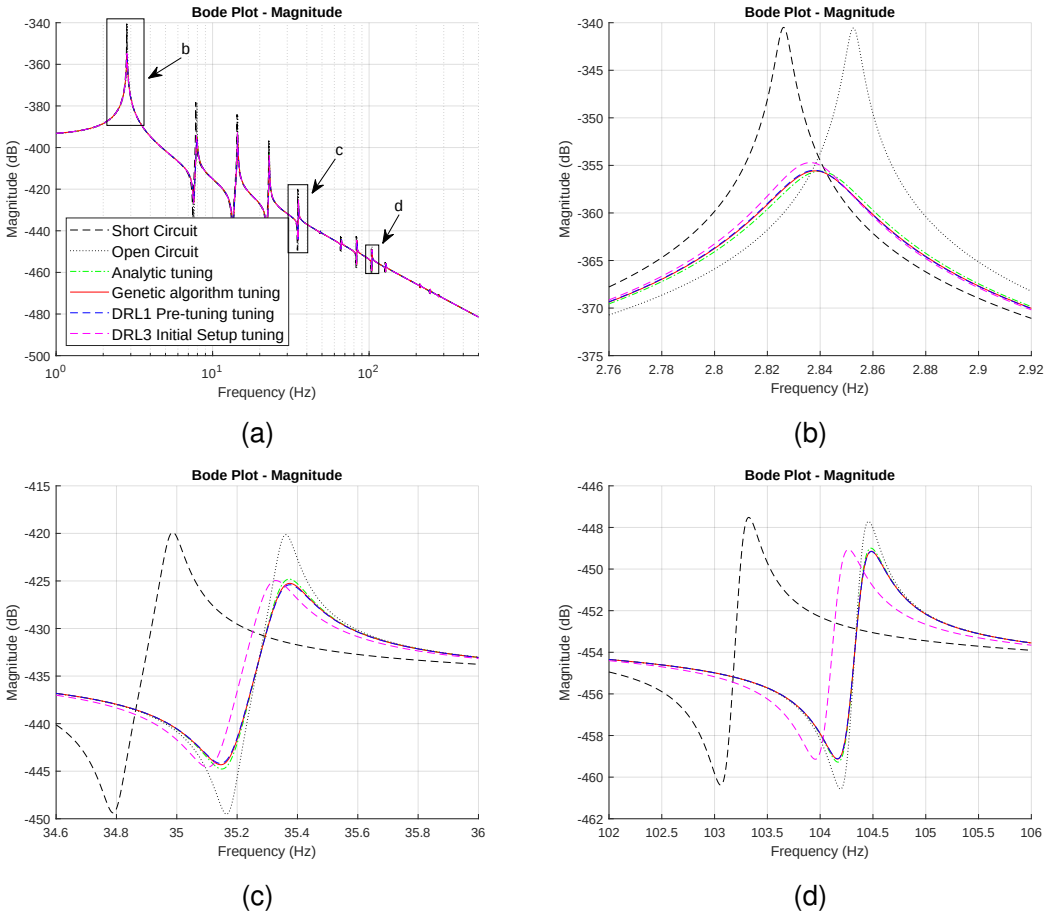


Figure 6.16: Bode from Ft_1 to w_1 (a), with zoom on the first mode (b), the fifth (c) and the ninth (d) bending mode: R shunt - Distinct case

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Table 6.6: Shunt values and reward computed from the numerical model: R shunt

Tuning Method	Function Count	$log_{10}(R)$ MEAN	$log_{10}(R)$ STD	Reward
Short circuit	-	0	0	0
Identical Shunt				
Reference	-	5.7532	0	24.86
GA	3014	5.6942	0	25.07
DRL0	500	5.6970	0	25.07
Distinct Shunt				
GA	34410	5.6933	0.0075	25.07
DRL1 (from Pre-Training)	35000	5.6923	0.0564	24.95
DRL3 (from Initial Setup)	35000	4.5600	2.2492	22.81

6.2.3.2 Control with R-C_{neg} shunt on the model

Optimisation algorithms for tuning a resistive-capacitive shunt are run using the same steps as for the resistive shunt (Figure 6.13). Figure 6.17 monitors the training and Table 6.7 shows the tuning results of the different methods.
Considering identical shunt tuning with two parameters, both DRL0 and Genetic Algorithm outperform the analytical method in terms of reward. Although the parameters are still close (Table 6.7) and the system response with the three methods looks similar as shown in Figure 6.18. The genetic algorithm requires 4525 estimations to reach the optimal found value whereas DRL0 needed only 200 episodes to reach a similar result. DRL0 is not trained on more episodes to avoid over-fitting phenomenon.
With 40 parameters to tune for the distinct case, the genetic algorithm method requires more than 150,000 estimates and was stopped due to time constraints. With pre-training of DRL0, DRL1 succeeds in reaching a solution, whereas without pre-training the optimisation fails with DRL3. The training of DRL1 and DRL3 can be seen in figures 6.17b and 6.17c. At the beginning of the training, the DRL1 reward is below the GA reward, but close. As the episodes progress, the reward increases until the average reward (over 10 episodes) exceeds that of the GA identical shunt case. In figure 6.17d, DRL1 training continues as DRL2 to observe an increase in reward above the identical case. Looking at the system response in Figure 6.19, the optimisations obtained from the different methods are very similar. The reward is sensitive to damping variations, so a reward delta of 1 is a neglected damping variation. DRL2 reaches a higher reward by spreading resistive and shunt values from the identical shunt case as indicated by the computed standard deviations.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

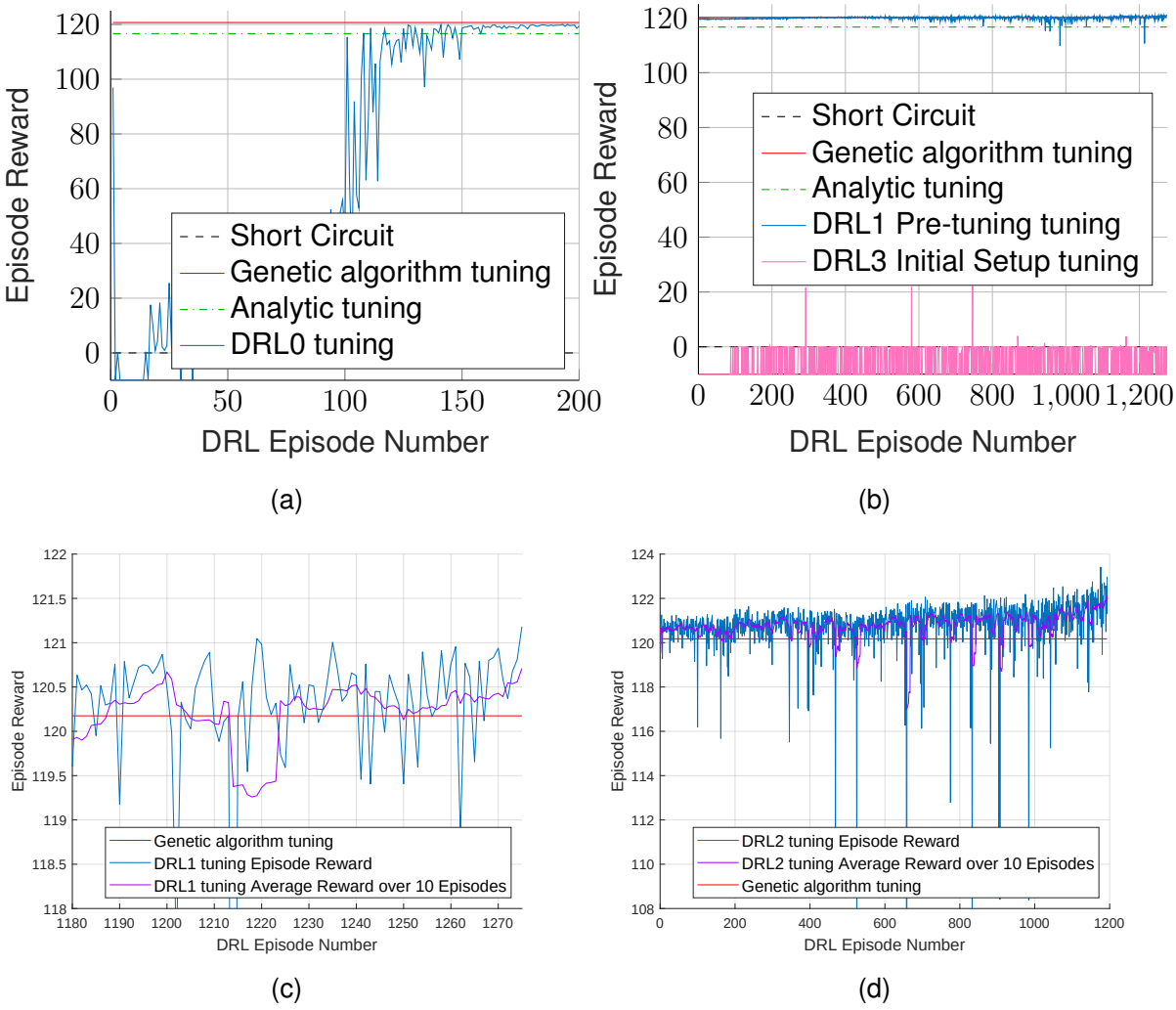


Figure 6.17: Training DRL0 (a), DRL1 and DRL3 (b), DRL1 with a zoom between episodes 1180 and 1270 (c) DRL2 (d): RC_{neg}

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

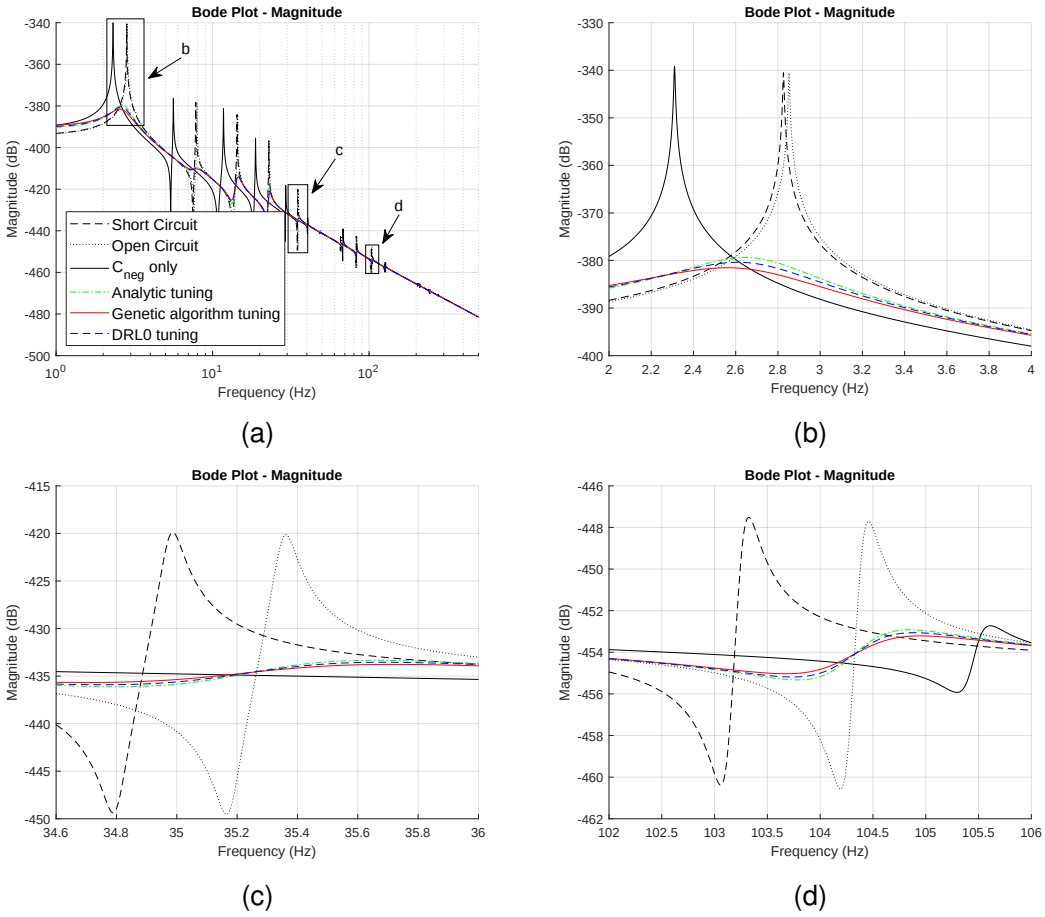


Figure 6.18: Bode from Ft_1 to w_1 (a), with zoom on the first mode (b), the fifth (c) and the ninth (d) bending mode: RC_{neg} shunt - Identical case

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

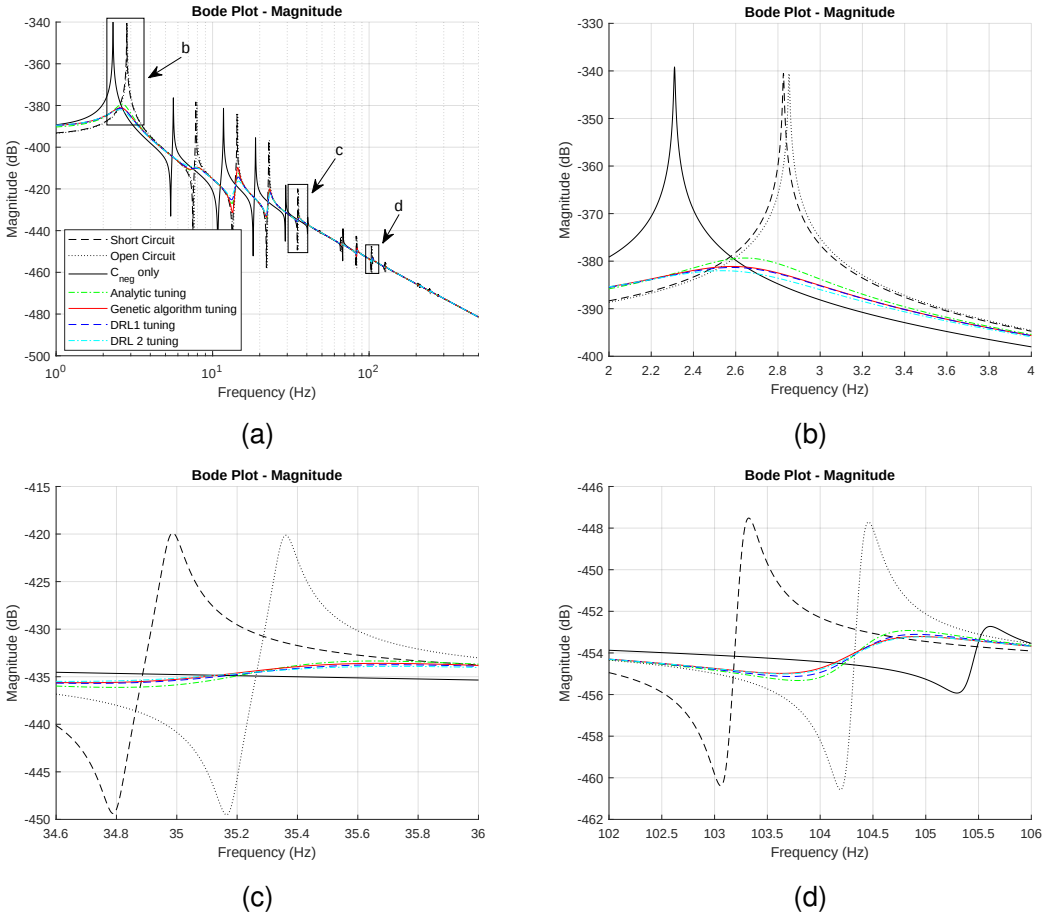


Figure 6.19: Bode from Ft_1 to w_1 (a), with zoom on the first mode (b), the fifth (c) and the ninth (d) bending mode: RC_{neg} shunt - Distinct case

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Table 6.7: Shunt values and reward computed from the numerical model: RC_{neg}

Tuning Method	Function Count	$log_{10}(R)$ MEAN	$log_{10}(R)$ STD	$log_{10}(-C_{neg})$ MEAN	$log_{10}(-C_{neg})$ STD	Reward
Short circuit	-	-	-	-	-	0
Identical Shunt						
Reference	-	4.5690	0	-6.9652	0	116.64
GA	4565	4.4465	0	-6.9704	0	120.68
DRL0	200	4.5137	0	-6.9678	0	119.42
Distinct Shunt						
GA	157944	5.2071	0.9896	-6.9507	0.1647	120.1746
DRL1 (from Pre-Training)	1275	4.4691	0.0955	-6.9698	0.0036	121.1669
DRL2 (from Pre-Training)	2468	4.4130	0.2076	-6.9696	0.0055	123.0682
DRL3 (from Initial Setup)	FAILED					-10

.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

6.3 Discussion and Conclusion

In this chapter, the DRL method successfully damped the vibrations of a smart beam by tuning a large number of parameters (40).
A finite element model of the beam was previously adapted to the experimental test bench. From this, a state space modal model is built to be used as a reduced numerical environment. Optimization methods are performed on this environment to tune transducers shunt parameters. The results presented above underline the following statements:

- The use of a resistive shunt with negative capacitance is more efficient in damping vibrations than a resistive shunt alone. The use of a negative capacitance also dampens vibration over a wider range of frequencies. These results were expected from the literature.
- DRL succeeds in tuning various shunt parameters to damp system vibrations, being as efficient as the analytical formulation for the resistive case and even better for the negative capacitance case.
- In this example application, identical and different shunt configurations result in the same attenuation performance. This is explained by the dominance of the first bending mode of the structure in the target frequency band, induced by the free boundary conditions.
- The training time of the DRL method is reduced by pre-training the NN. The pre-training is performed by reducing the number of actions to be performed on the environment. This technique is inspired by curriculum learning [6]. This machine learning method trains a model on examples with increasing problem complexity with episodes. In this way, the NN learns basic concepts before being enriched with more nuanced information. It helps to get better and faster results with convergence to the local optimum. In this way, DRL was able to find a solution for the $R-C_{neg}$ shunt.
- DRL converges faster than the Genetic Algorithm in this case using fewer function estimates.
- The model damping is not adjusted to match experimental measurements. However, it does not affect the optimal tuning result because it is focused on the first bending mode.

These experiments have shown that the proposed method using DRL is a reliable tool that can be used in more complex configurations for which tuning references are not known.

To go further, the DRL should be tested in the case where multiple modes are to be optimized. There the optimal values should be different depending on the transducer position on the beam. The learning curriculum could introduce a constraint of one mode to be optimized and then increase the number of target modes to ensure efficient tuning. The damping performance should also be tested on the experimental test bench.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Chapter 7

Conclusions

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

7.1 Main findings

This thesis aimed at proposing a novel tuning strategy for distributed control systems within the field of smart structures, ensuring stability and efficiency. It uses deep reinforcement learning (DRL) methods on classical control laws, to tune parameters for derivative, lead, PPF controllers, R-shunt, and RC negative shunt configurations.

The key findings highlight the efficiency of DRL for both scenarios of low and high complexity and with a small to large number of tunable parameters. However, challenges were observed due to the inherent randomness in initializing neural network parameters and training processes, which could lead to divergence. For larger-scale problems with more parameters, using pre-trained DRL models proved effective in avoiding divergence and accelerating the learning process. The approach demonstrated identical or even superior efficiency compared to traditional optimization methods such as analytic, simplex and genetic algorithms. This demonstrates the ability of DRL methods to handle diverse optimization criteria in both frequency and time domains, including non-continuous functions constrained by stability requirements.

The implications of this research are promising. The memory capabilities of neural networks allow for adaptability with minimal computational effort when environmental conditions change, offering a significant advantage over other optimization methods. This adaptability is especially relevant in real-world applications where system dynamics may vary.

Also, certain limitations were identified. The training process relies on interacting with a defined environment. In this thesis, a computational model was employed to ensure faster iterations, but discrepancies arose near stability limits due to model uncertainties. While the experimental test bench could be employed for DRL training, it demands robust backup systems to manage extended training durations and potential failures. In addition, the training process remains time-consuming (2 to 48 hours CPU) when working with a large number of values to optimize. However, other reinforcement learning strategies exist that are more efficient for large data dimensions, but these are not explored in this work. The development of the field of artificial intelligence has come up with new algorithms, but also with faster computing hardware such as GPUs.

7.2 Main results

Through the literature, Chapter 2 highlights the importance of considering transducers in the design process of a smart structure to provide maximum coupling coefficient, to transfer energy from the mechanical to the electrical domain. Classical active and passive control methods have been presented and show an effective ability to damp the vibration of a smart structure. However, the analytic tuning strategies often require additional manual tuning and expertise to avoid instabilities. These methods become impossible to use when a large number of values need to be tuned. In recent years, new strategies have been developed in the field of artificial intelligence, inspired by the human brain. Artificial neural networks are able to memorise patterns and handle large amounts of data. Many optimization algorithms have been developed and some applications have been implemented on smart structures to perform robust vibration control. Using these methods to mitigate vibration in smart structures is still emerging, especially in MIMO cases with centralized and non centralized strategies.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

In this thesis, it is chosen to use the reinforcement learning strategy using NN (called deep reinforcement learning - DRL) as an optimization algorithm to tune control laws without considering a database, but a defined environment. The proposed methodology is presented in Chapter 3 on a 1-dimensional smart cantilever beam with two piezoelectric collocated transducers. It efficiently tunes a resistive shunt with the aim of damping the first mode of vibration.

The method is then applied to the experimental test-bench configuration of a smart cantilever beam in Chapter 4. The structure is first designed using a finite element model to maximize the coupling coefficient and identify the modal compositions. A transfer function model is then built from the measurements using pole-zero identification. The DRL is then trained on the model of the beam to avoid damaging the experimental elements and allow faster convergence by reducing the number of measurements and the communication time between the DRL and the structure environment. A preliminary approach shows the effectiveness of DRL in tuning a derivative controller by considering only stable values. The methods are then used to efficiently tune a lead controller with three parameters. Tuned controllers are adapted to the input disturbance, considering noise and chirp signals, and take into account stability constraints introduced by the reward. The various controllers were tested on the test bench structure and successfully damped vibration in the selected frequency band including three bending modes and avoiding spillover.

In Chapter 5, the method is applied to an smart MIMO structure with three pairs of collocated transducers controlled independently. As in the previous chapter, the structure is first modelled using FEM. The third and fifth bending modes are the main contributors to the frequency response. It is therefore decided to use distributed feedback PPF controllers to damp the third bending mode of the structure. A transfer function model obtained from measurement identification is used as an environment to train optimization algorithms have also been introduced in the closed loop, leading to differences between the model and the test bench. A statistical approach is used to compare DRL and Simplex optimization strategies with a manually tuned reference controller. The results show the efficiency of DRL in tuning different types of controllers that lead to the same reward values. It confirms the possibility of reaching a local maximum or a sub-optimal solution. DRLs are more reliable than Simplex in providing a stable solution, with 52% of unstable solutions found for Simplex and 25% for DRL given the same initialization values. DRL also required fewer calculations to find a solution than Simplex (5000 function estimates for DRL and 1 million for Simplex per training). The robustness of the DRL method is demonstrated by its ability to adapt to variations in the objective function by changing the target frequency band or improving stability criteria.

In Chapter 6, the method is developed on a large MIMO smart structure with 20 to 40 parameters to be tuned with the aim of damping vibrations using shunted piezoelectric transducers. A finite element model based on the experimental setup is used to train the algorithms. DRL succeeds in tuning resistive and negative capacitance shunts and even outperforms analytical tuning. The damping is mostly focused on the first bending mode, as it is the most influential in the vibration response according to the free boundary conditions. In this application example, the DRL converges faster than genetic algorithms. Using a pre-trained DRL increases the convergence speed and accuracy compared to a random initialization of NNs. The use of a pre-training avoids the training failing during the process of finding a stable solution, especially with the negative capacitance shunt.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

7.3 Implementing deep reinforcement learning for control optimization: Main steps

Although AI tools are now widely used in everyday life, they have mainly been developed to interact with numerical environments. The challenge of this thesis was to adapt these tools to an experimental case of vibration control. This section presents the major questions that have arisen in the proposed method development and the corresponding proposed answers. It suggests practical guidelines for adapting the method to any vibration control system. Other online resources are available for more general applications.

7.3.1 Preliminary questions

The problem to be solved must first be defined by means of preliminary questions. The answers to these questions are given using this thesis example: vibration control optimization on a smart structure with transducers.

a. What needs to be optimized ?

The first question is trivial and common to all optimization problems in vibration control of structures:

- It suggests to identify an objective: minimize vibrations in a specified frequency range.
- This goal must be quantifiable : a measurable data is used to quantify the vibration response of the system through the sensors such as the piezoelectric transducers.
- The objective should be achievable: trying to control the modes of a structure that is already damped is pointless.
- The variable to be optimized should be defined: changes in the behaviour of the structure are achieved using actuators such as piezoelectric transducers. Feedback control laws are applied to the structure with a defined set of parameters to be adjusted.

The use of an achievable target called controllability is crucial to determine the success criteria for the control.

b. Are AI with neural networks tools relevant compared to traditional optimization methods like simplex or genetic algorithms ?

- No, especially when a limited number of interdependent values need to be estimated.
- Yes, AI and neural network tools are capable, thanks to their non-linear construction, of establishing connections by considering a large number of data.

Having an a priori idea of the expected result is useful for defining the boundaries of the study and avoiding unrelevant areas of research. Obviously, algorithms should be tested from the simplest to the most complex to identify the effectiveness of different reference optimization methods. The thesis used a DRL methods with basic NN but in a very high-dimensional case, the use of more sophisticated NNs such as convolutional neural networks (CNNs) could be considered.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

c. **How can AI be used in the context of smart structures?**

As can be seen from the literature (Chapter 2, Part 3), there are several possible applications of AI in the context of vibration control of structures:

- Artificial neural network controller: Neural networks are used directly as controllers, processing time-domain data to adjust the setpoint over time. This method requires a special experimental test-bench setup to minimise ADC and DAC transfer delays and achieve stable real-time control. In most publications, the learning phase requires a digital model of the structure.
- Structural design methods using AI as an optimizer is in line with the vision proposed in this thesis. The AI is an automation capable of adjusting control law parameters in the face of environmental variability and uncertainty.

d. **How can the AI be set up to control a structure?**

The main steps are as follow:

- Selection of the AI algorithm
- Define the elements of the algorithm: environment, observations, actions reward function and agent dimensions.
- Training the algorithm
- Critical analysis of results to avoid inconsistent results

The specific details of each element are described in the following section.

7.3.2 Selection of the AI algorithm

As seen in Chapter 2, Part 2, artificial intelligence includes a wide range of existing algorithms. This section focus only on machine learning algorithm using neural networks.

a. **Is there a database available ?**

- Yes, machine learning algorithms can be used with labelled or unlabelled databases. Issues around the amount of data required to condition the problem are not studied in this thesis.
- No, the use of a reinforcement learning method can be explored. The algorithm learns directly on the structure using a trial-and-error process, generating only the amount of data needed for the given optimization.

b. **Which reinforcement learning algorithm should be used ?**

The choice of reinforcement learning algorithm, also called agent, depends on the type of data to be estimated, a set of finite values or continuous data in an interval [2]. Each method has its advantages and disadvantages in terms of training speed, parallelization, memory efficiency, robustness or ease of hyperparameter tuning. This thesis focuses on the TRPO algorithm because it guarantees monotonous improvements and avoids falling into the unstable domain.

c. **What programming language should be used ?**

The two main languages to be considered in the context of AI for vibrating structures are currently Matlab and Python.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

- Matlab provides reinforcement learning and system identification and control libraries that are quick and easy to learn, with built-in help and tutorials. The drawback is that the internal functions of the library are a black boxes and not all functions can be customized. Also, the software is not open source and the latest RL algorithms are not always available.
- Python provides open access libraries created by users, often linked to published work, and offers a fully customization. The disadvantage for beginners is the number of options and choices to find a trustworthy library.

d. How to set the AI algorithm hyperparameters ?

Each RL algorithm comes with a set of hyperparameters that affect the training (Chapter 3, Part 1.4.2). The default settings work in most cases. They can be adjusted by calibration, but it is advisable to check the definitions of the other elements first. A parametric influence study can be performed to adjust the training in terms of learning speed or exploration.

7.3.3 Element Definitions

Reinforcement learning algorithms are based on a Markov decision process and include an environment, observations, an agent, actions and a reward. The environment is represented by a state, which reflects the conditions sent to the reinforcement learning (RL) agent. The agent uses these observations to select actions to perform on the environment, resulting in changes to its state. Depending on the effect of these actions, the environment generates a reward that represents their impact. The definition of each element is given according to the context of the thesis.

a. Environment

The environment is the structure to control. The use of a model helps to speed up training because the interface between the environment and the RL algorithm is put in a numerical framework. It also offers the possibility of not constraining the RL algorithm too much, which can try unstable control laws without damaging a testbench. The environment may be fixed, as studied in the thesis, or it may vary, for example, with a structure subjected to variable loading constraints.

b. Observations

The observations depend on the information that can be extracted from the environment. They correspond to data collected from sensors available on the structure. They need to be formatted and scaled in order to be used by the algorithm, preferably by normalizing them to avoid order-of-magnitude differences between values. To go further, the rules of observability and controllability used in automatic control should be transposed to establish a link between the complexity of the system to be studied and the minimum number of observations required to control it.

c. Agent

This thesis uses DRL through the TRPO and PPO methods, where an agent is divided into two neural networks: an actor that chooses the action and a critic that estimates the expected reward from observations of the environment. Questions arise about the design of these neural networks composed of input, output and hidden layers with activation functions.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

- **What size neural network is needed ?**

The number of neurons in the input layer corresponds to the number of observations. The number of neurons in the output layer is related to the number of actions. So this question is about the number of hidden layers and the number of neurons per hidden layer. The number of neurons in the hidden layers depends on the input and output neurons, but also on the complexity of the input-output function to be modeled. There aren't really any strict laws, but sizing is done more by trial error calculation (Chapter 3, Part 3.3). Some indicative guidelines can be found online in the free access training about AI, but are not mentioned here. Make sure that the size of the network is not too large, so that there is a trade-off between computational cost and data storage. Researchers are investigating the NN size influence on the learning process through episodes [33]. These studies examine the effects of various design elements, including the choice of activation function, the number of layers, the number of neurons per layer, the input metrics, and the training parameters.

- **Which activation function should be set ?**

The choice of activation function depends on the nature of the problem. Each function has its own subtleties, but in most cases the use of the Relu function in the hidden layers is appropriate because it adds nonlinearity to the network. Note that each activation function has an associated function for weight and bias initialization.

d. Actions

The definition of the actions depends on the adjustable parameters of the environment. In this work, the actions are defined according to the adjustable parameters of the control laws. To improve the training speed, the actions should be ranged in an interval. To avoid NN divergence, the actions should be scaled with an appropriate output function or with the same order of magnitude. For example, in Chapters 3 and 6 the resistance and capacity values were calculated using the power of 10 of the chosen action values, so the range of research was between -8 and 8, while being between 10^{-8} and 10^{8}.

e. Reward

The reward is the key point in the learning process as it represents the function to be optimized. It is advisable to set the reward values to the same order of magnitude as the network inputs and outputs. It is also important to avoid having two identical reward values for two different action values. The reward definition should limit the scope of the system as much as possible to avoid leading the algorithm to an unrealistic result. A conditional reward can be used, as in this thesis, to introduce a stability constraint. The reward function need not be convex over the whole range of data, but note that training may lead to a local optimum. Expertise in the application domain, such as mechanics in this thesis, is essential to the design of an efficient reward function.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

7.3.4 Training

The training process is controlled by the agent algorithm. Here are some practical guidelines to facilitate training success based on issues observed during the thesis.

a. Define a training end condition

The end of training condition is a value depending on the control objective, previously targeted and quantified. When this value is reached, training stops. If training is continued, it should be done step by step, because if the agent learns too much, the algorithm can cause a drop in performance.

b. Training by stages

Step-by-step training includes several elements:

- **Training by segmenting the learning:** By saving the network after a certain number of episodes, it is possible to monitor learning and avoid drop-outs, which would mean starting the training all over again. Moreover, this intermediate storage would make possible to observe the parallel evolution of 2 identical networks, for example by changing the reward function.
- **Training by gradually increasing the difficulty:** For example, in Chapter 6, training was done with a small number of parameters (one or two) to tune before increasing the number of parameters (20 or 40).
- **Training by trying several network initializations:** In Chapter 5 it was observed that the initialization of NN affects the speed and the convergence of the algorithm.

c. Play with random seed

The initialization of NN is done in this thesis with random functions. However, there is no real randomness in computers and this effect is controlled by a so-called 'seed'. In this work, as in most applications, the seed is initialized by the user and can therefore be modified. To achieve pseudo-randomness, a combination of seed functions was used in Chapter 4, with the cluster clock used to initialize the default seed.

d. Using a computing cluster

For small-scale calculations that take a few minutes to run, a personal computer is sufficient. However, even if the performance of a cluster is equivalent, the advantage is that a large number of simulations can be run simultaneously, for example to test different seeds on the same system.

7.3.5 Results

As with any model, AI results should be analysed and criticized by an expert, especially when applied to real-world physical systems. Sometimes the method produces adherent and even unrealistic results (Chapter 5, Part 3). This happens because the training process exploits the limits of the optimization function without any other prior knowledge. So it's important to design the reward properly, without introducing unnecessary bias.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

7.4 Ecological Cost

Depending on the problem to be solved, AI requires a variable amount of data and a corresponding amount of training time. For example, according to the Stanford Artificial Intelligence Index report [88] and David Patterson et al.[107], the initial training of chat GPT3 required 552 tons of C02 equivalent (C02e). In this case, this training phase consumes less energy than the amount needed to store the model in data centers, which depends on the location and time of storage. The energy consumption due to user queries depends on the type of response and can vary from 5g C02e for text to 200g of C02e for image generation [82]. In this thesis, the AI models used are relatively simple and single-task, so this entire work lasting three years, required 167.5 kg of C02 equivalent for digital equipment, including construction and use. Calculation details are given in Appendix F. To compare, flying from Paris to New York is equivalent to emitting approximately 887kg of CO2, considering a distance of 5836 km [5].

7.5 Future works: Going further with mechanical applications

7.5.1 Short term: Mechanical tools

As a short term improvement to this work, the model identification should be improved to better account for controller effects on the system. Furthermore, the developed method can be applied to larger scale distribution systems.

7.5.2 Mid-term: Use full capacities of AI tools

With the proposed method, the optimization can consider multiple states of the structures, such as different perturbation conditions, different boundary conditions or non-linearity. Considering smart structures, future research should explore the application of DRL in evolving environments. In this thesis, the systems were considered with only two states: the initial and the controlled state. The DRL methods offer the possibility to consider multiple dependent states and to optimize them all simultaneously with the same agent as introduced in Chapter 3 part 5.3.2. This strategy offers the possibility of creating embedded autonomous vibration controllers that can adapt to environmental variations and structural changes, and perform multiple intelligent functions simultaneously, such as wave focusing, damping or wave trapping.

7.5.3 Long-term: Adapt AI tools for mechanical applications

In addition to research aimed at developing new learning algorithms, it would be necessary to understand the links between the tools already developed, such as neural networks, and the complexity of the problems to be solved. The creation of design rules based on the study of the learning process would enable these algorithms to be used more effectively in connection with frugal AI. Investigating the relationship between neural network size and problem complexity, quantifying adaptability in response to environmental variations, and implementing the approach on more complex structures, such as composites, are key

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

areas for further study. The improvement of neural network training could be investigated by adding physics to the update functions, or by correlating the improvement of training behaviour with physical behaviour.

7.5.4 Last words

Finally, this thesis explores an efficient proven approaches to optimize the control of adaptive distributive vibratory systems. It offers a new perspective of optimization where the classical approach has failed. It bridges the fields of artificial intelligence and smart structures, offering new and creative perspectives to advance the state of the art in smart control systems.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Bibliography

[1] S. Ahmadi et al. “Electric-Responsive Materials: Properties, Design, and Applications”. In: *Stimuli-Responsive Materials for Biomedical Applications*. American Chemical Society, Mar. 2023, pp. 31–52. ISBN: 9780841297234. DOI: 10.1021/bk-2023-1436.ch002.

[2] F. AlMahamid and K. Grolinger. “Reinforcement Learning Algorithms: An Overview and Classification”. In: *2021 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)* (2021), pp. 1–7. URL: https://api.semanticscholar.org/CorpusID:240003074.

[3] G. L. Anderson, A. Crowson, and J. Chandra. “Introduction to Smart Structures”. In: *Intelligent Structural Systems*. Ed. by H. S. Tzou and G. L. Anderson. Dordrecht: Springer Netherlands, 1992, pp. 1–8. ISBN: 978-94-017-1903-2. DOI: 10.1007/978-94-017-1903-2_1. URL: https://doi.org/10.1007/978-94-017-1903-2_1.

[4] I. Asimov. *Runaround*. Ed. by A. S. Fiction. Street and Smith, 1942.

[5] *Avion long courrier | Impact CO2*. fr. URL: https://impactco2.fr/outils/transport/avion-longcourrier (visited on 12/09/2024).

[6] Y. Bengio et al. “Curriculum learning”. In: *Proceedings of the 26th Annual International Conference on Machine Learning*. ICML ’09. Montreal, Quebec, Canada: Association for Computing Machinery, 2009, pp. 41–48. ISBN: 9781605585161. DOI: 10.1145/1553374.1553380. URL: https://doi.org/10.1145/1553374.1553380.

[7] M. Berardengo et al. “Vibration control with piezoelectric elements: The indirect measurement of the modal capacitance and coupling factor”. In: *Mechanical Systems and Signal Processing* 151 (Apr. 2021), p. 107350. ISSN: 0888-3270. DOI: 10.1016/j.ymssp.2020.107350. URL: https://www.sciencedirect.com/science/article/pii/S0888327020307366 (visited on 08/09/2024).

[8] M. Berardengo et al. “Piezoelectric resonant shuntenhancement by negativecapacitances: Optimisation,performance and resonancecancellation”. In: 29 (2018), pp. 2581–2606. ISSN: 1045-389X. DOI: 10.1177/1045389x18770874.

[9] F. Berthoud et al. *Estimation de l’empreinte carbone d’une heure.coeur de calcul.* Tech. rep. UGA - UniversitéGrenoble Alpes; CNRS; INP Grenoble; INRIA., 2020. DOI: hal-02549565v5. URL: https://hal.science/hal-02549565v5/document.

[10] K. Billon et al. “Vibration isolation and damping using a piezoelectric flextensional suspension with a negative capacitance shunt”. In: *Mechanical Systems and Signal Processing* 140 (June 2020), p. 106696. ISSN: 0888-3270. DOI: 10.1016/j.ymssp.2020.106696. URL: https://www.sciencedirect.com/science/article/pii/S0888327020300820 (visited on 12/04/2024).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

[11] R. P. Borase et al. “A review of PID control, tuning methods and applications”. In: *International Journal of Dynamics and Control* 9.2 (July 2020), pp. 818–827. DOI: 10.1007/s40435-020-00665-4.

[12] A. Bryson. “Optimal control-1950 to 1985”. In: *IEEE Control Systems Magazine* 16.3 (June 1996), pp. 26–33. ISSN: 1941-000X. DOI: 10.1109/37.506395. URL: https://ieeexplore.ieee.org/document/506395 (visited on 07/15/2024).

[13] D. Cardon, J.-P. Cointet, and A. Mazières. “La revanche des neurones: L’invention des machines inductives et la controverse de l’intelligence artificielle”. In: *Réseaux* n°211.5 (Nov. 2018), pp. 173–220. ISSN: 0751-7971. DOI: 10.3917/res.211.0173.

[14] A. Cauchy. “Méthode générale pour la résolution des systèmes d’équations simultanées”. In: *Comptes Rendus de l’Académie des Sciences de Paris* t.25 (1847), pp. 536–538. URL: https://gallica.bnf.fr/ark:/12148/bpt6k2982c/f540.

[15] Y. Chauvin and D. E. Rumelhart, eds. *Backpropagation: Theory, Architectures, and Applications*. New York: Psychology Press, Feb. 1995. ISBN: 9780203763247. DOI: 10.4324/9780203763247.

[16] J. Y. Cheng and T. Mailund. “Ancestral population genomics using coalescence hidden Markov models and heuristic optimisation algorithms”. In: *Computational Biology and Chemistry*. 13th Asia Pacifi c Bioinformatics Conference, HsinChu, Taiwan, 21-23 January 2015 57 (Aug. 2015), pp. 80–92. ISSN: 1476-9271. DOI: 10.1016/j.compbiolchem.2015.02.001. URL: https://www.sciencedirect.com/science/article/pii/S1476927115000183 (visited on 08/11/2024).

[17] K. L. Chien, J. A. Hrones, and J. B. Reswick. “On the Automatic Control of Generalized Passive Systems”. In: *Transactions of the American Society of Mechanical Engineers* 74.2 (July 1952), pp. 175–183. ISSN: 0097-6822. DOI: 10.1115/1.4015724. URL: https://doi.org/10.1115/1.4015724 (visited on 12/03/2024).

[18] B. Chomette. *MIMO Modal Parameters Identification in Frequency Domain*. 2024. URL: https://www.mathworks.com/matlabcentral/fileexchange/82380-mimo-modal-parameters-identification-in-frequency-domain.

[19] G. H. Cohen and G. A. Coon. “Theoretical Consideration of Retarded Control”. In: *Transactions of the American Society of Mechanical Engineers* 75.5 (July 1953), pp. 827–834. ISSN: 0097-6822. DOI: 10.1115/1.4015451. URL: https://doi.org/10.1115/1.4015451 (visited on 12/03/2024).

[20] M. Collet and K. A. Cunefare. “Modal Synthesis and Dynamical Condensation Methods for Accurate Piezoelectric Systems Impedance Computation”. In: 19 (2008), pp. 1251–1269. ISSN: 1045-389X. DOI: 10.1177/1045389x07084956.

[21] M. Collet et al. “Vibroacoustic Energy Diffusion Optimization in Beams and Plates by Means of Distributed Shunted Piezoelectric Patches”. In: 2011, pp. 265–302. DOI: 10.1007/978-94-007-1703-9_10.

[22] B. Crepy et al. “The M4 adaptive unit for the E-ELT”. In: *1st AO4ELT conference - Adaptive Optics for Extremely Large Telescopes*. AO4ELT. EDP Sciences, 2010. DOI: 10.1051/ao4elt/201006001.

[23] H. Cudney, D. Inman, and Y. Oshman. “Distributed structural control using multilayered piezoelectric actuators”. In: (Apr. 1990). DOI: 10.2514/6.1990-1088.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

[24] P. Curie. *Oeuvres de Pierre Curie / publ. par les soins de la société française de physique*. Ed. by Paris. Bibliothèque nationale de France, 1984. URL: https://gallica.bnf.fr/ark:/12148/bpt6k2282p/f9.item.zoom.

[25] J. Doyle et al. “State-space solutions to standard H/sub 2/ and H/sub infinity / control problems”. In: *IEEE Transactions on Automatic Control* 34.8 (1989), pp. 831–847. DOI: 10.1109/9.29425.

[26] A. Erturk and D. J. Inman. *Piezoelectric Energy Harvesting*. Wiley, Mar. 2011, pp. 49–96. ISBN: 9781119991151. DOI: 10.1002/9781119991151.

[27] J. L. Fanson and T. K. Caughey. “Positive position feedback control for large space structures”. In: *AIAA Journal* 28.4 (Apr. 1990), pp. 717–724. ISSN: 1533-385X. DOI: 10.2514/3.10451.

[28] S. Fenik and L. Starek. “Optimal PPF Controller for Multimodal Vibration Suppression”. In: (Jan. 2008).

[29] R. A. Fisher. “The Use of Multiple Measurements in Taxonomic Problems”. en. In: *Annals of Eugenics* 7.2 (1936), pp. 179–188. ISSN: 2050-1439. DOI: 10.1111/j.1469-1809.1936.tb02137.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1469-1809.1936.tb02137.x (visited on 06/16/2023).

[30] P. Fleming and R. Purshouse. “Evolutionary algorithms in control systems engineering: a survey”. In: *Control Engineering Practice* 10.11 (Nov. 2002), pp. 1223–1241. ISSN: 0967-0661. DOI: 10.1016/s0967-0661(02)00081-3.

[31] G. Franklin, J. Powell, and A. Emami-Naeini. *Feedback Control of Dynamic Systems*. Ed. by Addison-Wesley. 1986.

[32] G. Friedland, M. M. Krell, and A. Metere. *A Practical Approach to Sizing Neural Networks*. Sept. 2018. DOI: 10.2172/1476219.

[33] G. Friedland, M. M. Krell, and A. Metere. *A Practical Approach to Sizing Neural Networks*. English. Tech. rep. Sept. 2018. DOI: 10.2172/1476219. URL: https://www.osti.gov/biblio/1476219 (visited on 06/19/2024).

[34] K. Fukushima. “Visual Feature Extraction by a Multilayered Network of Analog Threshold Elements”. In: *IEEE Transactions on Systems Science and Cybernetics* 5.4 (1969), pp. 322–333. DOI: 10.1109/tssc.1969.300225.

[35] C. Galeazzi et al. “Analysis and Control of Microvibrations on ARTEMIS Satellite”. In: *Journal of Intelligent Material Systems and Structures* 7.2 (Mar. 1996), pp. 216–226. DOI: 10.1177/1045389x9600700214.

[36] T. Gang and F. L. Lewis, eds. *Adaptive Control with Input Saturation Constraints*. Springer London, 2001, pp. 361–381. ISBN: 9781447136873. DOI: 10.1007/978-1-4471-3687-3_15.

[37] V. Giurgiutiu. “Actuators and Smart Structures”. In: ed. by S. Braun. Oxford: Elsevier, Jan. 2001, pp. 58–81. ISBN: 9780122270857. DOI: 10.1006/rwvb.2001.0197. URL: https://www.sciencedirect.com/science/article/pii/B0122270851001971 (visited on 12/06/2024).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

[38] C. J. Goh and T. K. Caugey. “On the stability problem caused by finite actuator dynamics in the collocated control of large space structures”. In: *International Journal of Control* 41.3 (Mar. 1985), pp. 787–802. ISSN: 0020-7179. DOI: 10.1080/00207185089611163. URL: https://doi.org/10.1080/00207185089611163 (visited on 08/05/2024).

[39] D. E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning*. en. Google-Books-ID: 2IIJAAAACAAJ. Addison-Wesley, 1989. ISBN: 9780201157673.

[40] A. Guerra-Langan, S. Araujo Estrada, and S. Windsor. “Reinforcement Learning to Control Lift Coefficient Using Distributed Sensors on a Wind Tunnel Model”. In: *AIAA SCITECH 2022 Forum*. American Institute of Aeronautics and Astronautics, Jan. 2022. DOI: 10.2514/6.2022-0966.

[41] E. Gulbahce and M. Celik. “Fuzzy logic aided PPF controller design to active vibration control of a flexible beam”. In: *Journal of Structural Engineering and Applied Mechanics* 4.3 (Sept. 2021), pp. 184–195. ISSN: 2630-5763. DOI: 10.31462/jseam.2021.03184195.

[42] J. A. Guzmán, G. Pizarro, and F. Núñez. “A Reinforcement Learning-Based Distributed Control Scheme for Cooperative Intersection Traffic Control”. In: *IEEE Access* 11 (2023), pp. 57037–57045. ISSN: 2169-3536. DOI: 10.1109/access.2023.3283218.

[43] N. W. Hagood and A. von Flotow. “Damping of Structural Vibrations with PiezoelectricMaterials and Passive Electrical Networks, Journal of Sound and Vibration”. In: 146 (1991), pp. 243–268. ISSN: 0022-460X. DOI: 10.1016/0022-460x(91)90762-9.

[44] J.-H. Han and I. Lee. “Optimal placement of piezoelectric sensors and actuators for vibration control of a composite plate using genetic algorithms”. In: *Smart Materials and Structures* 8.2 (1999), p. 257. URL: https://iopscience.iop.org/article/10.1088/0964-1726/8/2/012/meta (visited on 07/23/2024).

[45] S. Hassani, M. Mousavi, and Z. Sharif-Khodaei. “Chapter 14 - Smart bridge monitoring”. In: ed. by A. H. Alavi et al. Butterworth-Heinemann, Jan. 2022, pp. 343–372. ISBN: 9780128177846. DOI: 10.1016/B978-0-12-817784-6.00007-2. URL: https://www.sciencedirect.com/science/article/pii/B9780128177846000072 (visited on 12/06/2024).

[46] K. P. T. Haughn, C. Harvey, and D. J. Inman. “Deep learning reduces sensor requirements for gust rejection on a small uncrewed aerial vehicle morphing wing”. In: *Communications Engineering* 3.1 (Mar. 2024). ISSN: 2731-3395. DOI: 10.1038/s44172-024-00201-8.

[47] M. Heskins and J. E. Guillet. “Solution Properties of Poly(N-isopropylacrylamide)”. In: *Journal of Macromolecular Science: Part A - Chemistry* 2.8 (Dec. 1968), pp. 1441–1455. ISSN: 0022-233X. DOI: 10.1080/10601326808051910.

[48] A. S. Hoffman. ““Intelligent” Polymers in Medicine and Biotechnology”. In: *Artificial Organs* 19.5 (May 1995), pp. 458–467. ISSN: 1525-1594. DOI: 10.1111/j.1525-1594.1995.tb02359.x.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

[49] J. H. Holland. *Genetic Algorithms*. en. July 1992. DOI: http://dx.doi.org/10.1038/scientificamerican0792-66. URL: https://www.scientificamerican.com/article/genetic-algorithms/ (visited on 12/05/2024).

[50] G. Housner et al. “Structural Control: Past, Present, and Future”. In: *Journal of EngineeringMechanics* 123 (1997), pp. 897–971. ISSN: 0733-9399. DOI: 10.1061/(asce)0733-9399(1997)123:9(897).

[51] R. A. Howard. *Dynamic Programming and Markov Processes*. The Massachusetts Institute of Technology, 1960. ISBN: 978-0262080095. URL: https://www.gwern.net/docs/statistics/decision/1960-howard-dynamicprogrammingmarkovprocesses.pdf.

[52] J. Hu et al. “Cooperative Internet of UAVs: Distributed Trajectory Design by Multi-Agent Deep Reinforcement Learning”. In: *IEEE Transactions on Communications* 68.11 (Nov. 2020), pp. 6807–6821. ISSN: 1558-0857. DOI: 10.1109/tcomm.2020.3013599.

[53] S. Hurlebaus and L. Gaul. “Smart structure dynamics”. In: *Mechanical Systems and Signal Processing* 20.2 (Feb. 2006), pp. 255–281. ISSN: 0888-3270. DOI: 10.1016/j.ymssp.2005.08.025. URL: https://www.sciencedirect.com/science/article/pii/S0888327005001330 (visited on 12/06/2024).

[54] J. S. R. Jang, C. T. Sun, and E. Mizutani. *Neuro-Fuzzy and Soft Computing—A Computational Ap-proach to Learning and Machine Intelligence*. Ed. by E. Cliffs. Prentice-Hall, 1997. ISBN: 0-13-291066-3. URL: http://www.soukalfi.edu.sk/01_NeuroFuzzyApproach.pdf.

[55] T. Jaunet, R. Vuillemot, and C. Wolf. “Understanding Decisions and Memory in Deep Reinforcement Learning - 1909.02982”. In: *Computer Graphics Forum* 39 (2020), pp. 49–61. ISSN: 0167-7055. DOI: 10.1111/cgf.13962.

[56] D. Jeanmaire et al. “Chemical specificity in REDOX-responsive materials: the diverse effects of different Reactive Oxygen Species (ROS) on polysulfide nanoparticles”. In: *Polymer Chemistry* 5.4 (2014), p. 1393. ISSN: 1759-9962. DOI: 10.1039/c3py01475d.

[57] R. Jha and J. Rower. “Experimental investigation of activevibration control using neural networksand piezoelectric actuators”. In: *SMS* 11 (2002), pp. 115–121. ISSN: 0964-1726. DOI: 10.1088/0964-1726/11/1/313.

[58] M. M. Jovanović et al. “Experimental studies on active vibration control of a smart composite beam using a PID controller”. In: *Smart Materials and Structures* 22.11 (Oct. 2013). ISSN: 1361-665X. DOI: 10.1088/0964-1726/22/11/115038.

[59] W. B. S. Jr, J. S. Sirkis, and P. T. Gardiner. “Smart materials and structures: what are they?” en. In: *Smart Materials and Structures* 5.3 (June 1996), p. 247. ISSN: 0964-1726. DOI: 10.1088/0964-1726/5/3/002. URL: https://dx.doi.org/10.1088/0964-1726/5/3/002 (visited on 12/06/2024).

[60] S. Katoch, S. S. Chauhan, and V. Kumar. “A review on genetic algorithm: past, present, and future”. en. In: *Multimedia Tools and Applications* 80.5 (Feb. 2021), pp. 8091–8126. ISSN: 1573-7721. DOI: 10.1007/s11042-020-10139-6. URL: https://doi.org/10.1007/s11042-020-10139-6 (visited on 12/05/2024).

[61] A. Khalatbarisoltani. “Online control of an active seismic system via reinforcement learning”. In: 26 (2019), e2298. ISSN: 1545-2255. DOI: 10.1002/stc.2298.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

[62] S. Khizar et al. “Magnetic-Responsive Materials: Properties, Design, and Applications”. In: *Stimuli-Responsive Materials for Biomedical Applications*. American Chemical Society, Mar. 2023, pp. 53–79. ISBN: 9780841297234. DOI: 10.1021/bk-2023-1436.ch003.

[63] M. Khorasaninejad et al. “Metalenses at visible wavelengths: Diffraction-limited focusing and subwavelength resolution imaging”. In: *Science* 352.6290 (June 2016), pp. 1190–1194. ISSN: 1095-9203. DOI: 10.1126/science.aaf6644.

[64] S. Khot et al. “Active vibration control of cantilever beam by using PID based output feedback controller”. In: *Journal of Vibration and Control* 18.3 (July 2012), pp. 366–372. ISSN: 1741-2986. DOI: 10.1177/1077546311406307.

[65] G. Kocak, C. Tuncer, and V. Bütün. “pH-Responsive polymers”. In: *Polymer Chemistry* 8.1 (2017), pp. 144–176. ISSN: 1759-9962. DOI: 10.1039/c6py01872f.

[66] W. Koch et al. “Reinforcement Learning for UAV Attitude Control”. In: *ACM Transactions on Cyber-Physical Systems* 3.2 (Feb. 2019), pp. 1–21. ISSN: 2378-9638. DOI: 10.1145/3301273.

[67] P. V. Kokotovic. “The joy of feedback: nonlinear and adaptive”. In: *IEEE Control Systems* 12.3 (June 1992), pp. 7–17. ISSN: 1066-033X, 1941-000X. DOI: 10.1109/37.165507. URL: https://ieeexplore.ieee.org/document/165507/ (visited on 07/15/2024).

[68] B. Kommey et al. “A Reinforcement Learning Review: Past Acts, Present Facts and Future Prospects”. In: *IT Journal Research and Development* 8 (Feb. 2024), pp. 120–142. DOI: 10.25299/itjrd.2023.13474.

[69] R. Kumar. “Effective active vibration control of single link flexible manipulator with modified positive position feedback control in the presence of instrumentation phase lead/lag”. en. In: *Journal of Vibration and Control* 19.10 (July 2013), pp. 1538–1560. ISSN: 1077-5463. DOI: 10.1177/1077546312448409. URL: https://doi.org/10.1177/1077546312448409 (visited on 08/06/2024).

[70] S. M. Kusagur, G. Arunkumar, and T. C. Manjunath. “Modelling of smart intelligent materials with PZT & PVDF sensors/actuators to control the active vibrations of flexible aluminum mechanical cantilever beams using proportional integral derivative (PID) techniques”. In: *Materials Today: Proceedings*. International Conference on Newer Trends and Innovation in Mechanical Engineering: Materials Science 37 (Jan. 2021), pp. 2075–2082. ISSN: 2214-7853. DOI: 10.1016/j.matpr.2020.07.507. URL: https://www.sciencedirect.com/science/article/pii/S2214785320356170 (visited on 08/06/2024).

[71] M. Kwak and D. Sciulli. “Fuzzy-Logic based vibration suppression control experiment on active structures”. In: *Journal of Sound and Vibration* 191.1 (Mar. 1996), pp. 15–28. DOI: 10.1006/jsvi.1996.0104.

[72] J. C. Lagarias et al. “Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions”. In: *SIAM J. Optim.* 9 (1998), pp. 112–147. URL: https://api.semanticscholar.org/CorpusID:9245771.

[73] G.-S. Lee. “System identification and control of smart structures using neural networks”. In: *Elsevier Science Ltd* 38 (1996), pp. 269–276. ISSN: 0094-5765. DOI: 10.1016/0094-5765(96)00053-7.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

[74] K. J. Lee. “Chapter Seven - Architecture of neural processing unit for deep neural networks”. In: *Hardware Accelerator Systems for Artificial Intelligence and Machine Learning*. Ed. by S. Kim and G. C. Deka. Vol. 122. Advances in Computers. Elsevier, 2021, pp. 217–245. DOI: https://doi.org/10.1016/bs.adcom.2020.11.001.

[75] G. Li et al. “Analysis of smart beams with piezoelectric elements using impedance matrix and inverse Laplace transform”. In: *Smart Materials and Structures* 22 (Sept. 2013), p. 115001. DOI: 10.1088/0964-1726/22/11/115001.

[76] L. Li, G. Song, and J. Ou. “Adaptive fuzzy sliding mode based active vibration control of a smart beam with mass uncertainty”. In: *Structural Control and Health Monitoring* (2009). ISSN: 1545-2255. DOI: 10.1002/stc.356.

[77] Y. Liang et al. “Neural network modeling and single-neuron proportional–integral–derivative control for hysteresis in piezoelectric actuators”. In: *Measurement and Control* 52.9–10 (Oct. 2019), pp. 1362–1370. ISSN: 0020-2940. DOI: 10.1177/0020294019866846.

[78] C.-Y. Lin, Y.-H. Huang, and W.-T. Chen. “Multimodal suppression of vibration in smart flexible beam using piezoelectric electrode-based switching control”. In: *Mechatronics* 53 (Aug. 2018), pp. 152–167. ISSN: 0957-4158. DOI: 10.1016/j.mechatronics.2018.06.005.

[79] W. Lippmann. *Annales de chimie et de physique*. Ed. by C. (V. M. M. (Paris). Bibliothèque nationale de France, 1881. URL: https://gallica.bnf.fr/ark:/12148/bpt6k348640/f568.item.

[80] S. Livet et al. “Turbomachinery blades damping thanks to optimized shunted piezoelectric circuits”. In: (2008). ISSN: 0277-786X. DOI: 10.1117/12.776898.

[81] S. Livet et al. “Structural multi-modal damping by optimizing shunted piezoelectric transducers”. In: (2011), pp. 73–102. ISSN: 2642-2050. DOI: 10.13052/ejcm.20.73-102.

[82] A. S. Luccioni, Y. Jernite, and E. Strubell. “Power Hungry Processing: Watts Driving the Cost of AI Deployment?” In: (2023). DOI: 10.48550/ARXIV.2311.16863.

[83] S. Lustig and D. Elata. “Ambiguous definitions of the piezoelectric coupling factor”. en. In: *Journal of Intelligent Material Systems and Structures* 31.14 (Aug. 2020), pp. 1689–1696. ISSN: 1045-389X. DOI: 10.1177/1045389X20930104. URL: https://doi.org/10.1177/1045389X20930104 (visited on 12/02/2024).

[84] K. Maity and D. Mandal. “Chapter Seven - Piezoelectric polymers and composites for multifunctional materials”. In: *Advanced Lightweight Multifunctional Materials*. Ed. by P. Costa, C. M. Costa, and S. Lanceros-Mendez. Woodhead Publishing in Materials. Woodhead Publishing, 2021, pp. 239–282. ISBN: 978-0-12-818501-8. DOI: https://doi.org/10.1016/B978-0-12-818501-8.00001-9. URL: https://www.sciencedirect.com/science/article/pii/B9780128185018000019.

[85] V. Makoviychuk et al. *Isaac Gym: High Performance GPU-Based Physics Simulation For Robot Learning*. 2021. DOI: 10.48550/ARXIV.2108.10470.

[86] W. Marckwald. “Ueber Phototropie”. In: *Zeitschrift für Physikalische Chemie* 30U.1 (Oct. 1899), pp. 140–145. ISSN: 0942-9352. DOI: 10.1515/zpch-1899-3007.

[87] E. Markiewicz. “Commande distribuée : application au contrôle actif de vibrations”. Theses. Ecole Centrale de Lyon, Feb. 2024. URL: https://theses.hal.science/tel-04766022.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

[88] N. Maslej et al. *The AI Index 2024 Annual Repor*. Tech. rep. AI Index Steering Committee, Institute for Human-Centered AI, StanfordUniversity, Stanford, CA, Apr. 2024. URL: https://aiindex.stanford.edu/wp-content/uploads/2024/05/HAI_AI-Index-Report-2024.pdf.

[89] Y. Matanga. "PhD Thesis : Convergence Acceleration in Global Optimisation with Application to Control Systems". PhD thesis. Dec. 2022.

[90] F. Maugan et al. "Enhancement of energy harvesting usingacoustical-black-hole-inspired wave traps". In: *Smart Materials and Structures* 28.7 (May 2019), p. 075015. DOI: 10.1088/1361-665X/ab1f11.

[91] W. McCulloch and W. Pitts. "A Logical Calculus of the Ideas Immanent in Nervous Activity (1943)". In: *Ideas That Created the Future*. The MIT Press, Feb. 1943, pp. 79–88. DOI: 10.7551/mitpress/12274.003.0011.

[92] A. H. Meitzler et al. "IEEE Standard on Piezoelectricity". In: (1988). DOI: 10.1109/ieeestd.1988.79638.

[93] M. Minsky. "A Framework for Representing Knowledge". In: (1988), pp. 156–189. DOI: 10.1016/b978-1-4832-1446-7.50018-2. URL: https://web.media.mit.edu/~minsky/papers/Frames/frames.html.

[94] M. Minsky and S. A. Papert. *Perceptrons: An Introduction to Computational Geometry. An Introduction to Computational Geometry*. 1969. DOI: 10.7551/mitpress/11301.001.0001. URL: https://api.semanticscholar.org/CorpusID:5400596.

[95] J. Mitchell. "Comments on "Bode compensator design"". In: *IEEE Transactions on Automatic Control* 22.5 (Oct. 1977), pp. 869–870. ISSN: 1558-2523. DOI: 10.1109/tac.1977.1101590.

[96] V. Mnih et al. *Playing Atari with Deep Reinforcement Learning*. 2013. DOI: 10.48550/ARXIV.1312.5602.

[97] Mohit, D. Chhabra, and S. Kumar. "Active Vibration Control of the Smart Plate Using Artificial Neural Network Controller". In: *Advances in Aerospace Engineering* 2015 (2015), pp. 1–20. ISSN: 2356-6531. DOI: 10.1155/2015/137068.

[98] P. Monnier, M. Collet, and J. Piranda. "Definition of mechanical design parameters to optimize efficiency of integral force feedback". In: *STRUCTURAL CONTROL AND HEALTH MONITORING* 12.1 (Jan. 2005), pp. 65–89. ISSN: 1545-2255. DOI: 10.1002/stc.57.

[99] J. A. Nelder and R. Mead. "A Simplex Method for Function Minimization". In: *The Computer Journal* 7.4 (Jan. 1965), pp. 308–313. ISSN: 0010-4620. DOI: 10.1093/comjnl/7.4.308. URL: https://doi.org/10.1093/comjnl/7.4.308 (visited on 08/11/2024).

[100] "Line Search Methods". In: *Numerical Optimization*. Ed. by J. Nocedal and S. J. Wright. New York, NY: Springer New York, 1999, pp. 34–63. ISBN: 978-0-387-22742-9. DOI: 10.1007/0-387-22742-3_3.

[101] O. D. I. Nwokah. "Pseudo-Derivative Feedback Control". In: *1987 American Control Conference*. Vol. 134. June 1987, pp. 1811–1814. DOI: 10.23919/ACC.1987.4789606. URL: https://ieeexplore.ieee.org/document/4789606 (visited on 08/02/2024).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

[102] H. Nyquist. “Regeneration Theory”. In: *Bell System Technical Journal* 11.1 (Jan. 1932), pp. 126–147. ISSN: 0005-8580. DOI: 10.1002/j.1538-7305.1932.tb02344.x.

[103] T. Ogawa. *Piezoelectric Materials*. InTech, Aug. 2016. ISBN: 9789535125594. DOI: http://dx.doi.org/10.5772/61563.

[104] M. Onoe, H. F. Tiersten, and A. H. Meitzler. “Shift in the Location of Resonant Frequencies Caused by Large Electromechanical Coupling in Thickness-Mode Resonators”. In: *The Journal of the Acoustical Society of America* 35.1 (Jan. 1963), pp. 36–42. ISSN: 0001-4966. DOI: 10.1121/1.1918410. URL: https://doi.org/10.1121/1.1918410 (visited on 12/02/2024).

[105] J. Panda et al. “An iterative gradient descent-based reinforcement learning policy for active control of structural vibrations”. In: *Computers and Structures* 290 (Jan. 2024), p. 107183. ISSN: 0045-7949. DOI: https://doi.org/10.1016/j.compstruc.2023.107183.

[106] W. K. H. Panofsky, M. Phillips, and P. M. Morse. “Classical Electricity and Magnetism”. In: 9 (1956), pp. 22–24. ISSN: 0031-9228. DOI: 10.1063/1.3059993.

[107] D. Patterson et al. *Carbon Emissions and Large Neural Network Training*. 2021. DOI: 10.48550/ARXIV.2104.10350.

[108] D. Pisarski and L. Jankowski. “Reinforcement learning-based control to suppress the transient vibration of semi-active structures subjected to unknown harmonic excitation”. In: *Computer-Aided Civil and Infrastructure Engineering* 38.12 (Sept. 2022), pp. 1605–1621. ISSN: 1467-8667. DOI: DOI:10.1111/mice.12920.

[109] J. R. Platt. “Electrochromism, a Possible Change of Color Producible in Dyes by an Electric Field”. In: *The Journal of Chemical Physics* 34.3 (Mar. 1961), pp. 862–863. ISSN: 1089-7690. DOI: 10.1063/1.1731686.

[110] J.-C. Pomerol. *Les sytèmes experts*. Ed. by Hermes. 1988. ISBN: 2-86601-153-8.

[111] A. Preumont. *A. Vibration Control of Active Structures: An Introduction*. 4th Edition. Springer Science and Business Media LLC, 2018, pp. 143–168. DOI: 10.1023/a:1004398914135.

[112] A. Preumont. *Vibration Control of ActiveStructures*. 2018, pp. 199–226. DOI: 10.1007/978-3-319-72296-2_9.

[113] Z.-c. Qiu, G.-h. Chen, and X.-m. Zhang. “Reinforcement learning vibration control for a flexible hinged plate”. In: *Aerospace Science and Technology* 118 (Nov. 2021), p. 107056. ISSN: 1270-9638. DOI: 10.1016/j.ast.2021.107056.

[114] H. R. Rahmani. “Artificial Intelligence Approach for Seismic Control of Structures”. PhD thesis. Bauhaus-Universität Weimar, 2020.

[115] B. Robert and R. Vittal. “Optimal control of infinite-order smart composite structural systems using distributed sensors”. In: *Composites Engineering* 4.6 (Jan. 1994), pp. 577–589. DOI: 10.1016/0961-9526(94)90102-3.

[116] J. Rodriguez, M. Collet, and S. Chesné. “Active Vibration Control on a Smart Composite Structure Using Modal-Shaped Sliding Mode Control”. In: *Journal of Vibration and Acoustics* 144.2 (Feb. 2022). DOI: 10.1115/1.4053358.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

[117] J. M. Rodriguez-Fortun et al. “Flatness-Based Active Vibration Control for Piezoelectric Actuators”. In: *IEEE/ASME Transactions on Mechatronics* 18.1 (Feb. 2013), pp. 221–229. ISSN: 1941-014X. DOI: 10.1109/TMECH.2011.2166998. URL: https://ieeexplore.ieee.org/document/6030944 (visited on 12/03/2024).

[118] F. Rosenblatt. “The perceptron: A probabilistic model for information storage and organization in the brain.” en. In: *OpenAIRE - Explore* 65 (1958), pp. 386–408. ISSN: 1939-1471. DOI: 10.1037/h0042519. URL: https://explore.openaire.eu/search/publication?pid=10.1037%2Fh0042519 (visited on 08/29/2023).

[119] S. J. Rupitsch. *Piezoelectric Sensors and Actuators: Fundamentals and Applications*. Springer Berlin Heidelberg, 2019. ISBN: 9783662575345. DOI: 10.1007/978-3-662-57534-5.

[120] G. B. Scharmer et al. “The 85-electrode adaptive optics system of the Swedish 1-m Solar Telescope”. In: *Astronomy and amp; Astrophysics* 685 (May 2024), A32. ISSN: 1432-0746. DOI: 10.1051/0004-6361/201936005.

[121] J. Schulman et al. “Proximal Policy Optimization Algorithms”. In: *ArXiv* (2017). DOI: https://arxiv.org/abs/1707.06347.

[122] J. Schulman et al. “Trust Region Policy Optimization”. In: *Proceedings of the 32nd International Conference on Machine Learning*. Ed. by F. Bach and D. Blei. Vol. 37. Proceedings of Machine Learning Research. Lille, France: PMLR, July 2015, pp. 1889–1897.

[123] A. K. Shakya, G. Pillai, and S. Chakrabarty. “Reinforcement learning algorithms: A brief survey”. In: *Expert Systems with Applications* 231 (Nov. 2023), p. 120495. ISSN: 0957-4174. DOI: 10.1016/j.eswa.2023.120495. URL: https://www.sciencedirect.com/science/article/pii/S0957417423009971 (visited on 08/16/2024).

[124] M. Y. Shalaginov et al. “Reconfigurable all-dielectric metalens with diffraction-limited performance”. In: *Nature Communications* 12.1 (Feb. 2021). ISSN: 2041-1723. DOI: 10.1038/s41467-021-21440-9.

[125] S. Sherrit et al. “Comparison of the Mason and KLM equivalent circuits for piezoelectric resonators in the thickness mode”. In: *1999 IEEE Ultrasonics Symposium. Proceedings. International Symposium (Cat. No.99CH37027)*. Vol. 2. 1999, 921–926 vol.2. DOI: 10.1109/ULTSYM.1999.849139.

[126] H. Shi et al. “A deep reinforcement learning based distributed control strategy for connected automated vehicles in mixed traffic platoon”. In: *Transportation Research Part C: Emerging Technologies* 148 (Mar. 2023), p. 104019. ISSN: 0968-090X. DOI: 10.1016/j.trc.2023.104019.

[127] F. G. Shinskey. *Process control systems: application, design and tuning*. EN. 4th ed. New York: McGraw-Hill, 1996. ISBN: 9780070571013.

[128] D. Silver et al. “Mastering the game of Go with deep neural networks and tree search”. en. In: *Nature* 529.7587 (Jan. 2016), pp. 484–489. ISSN: 1476-4687. DOI: 10.1038/nature16961. URL: https://www.nature.com/articles/nature16961 (visited on 09/05/2024).

[129] C. P. Smyser and K. Chandrashekhara. “Robust vibration control of composite-beams using piezoelectric devices and neural networks”. In: 6 (1997), pp. 178–189. ISSN: 0964-1726. DOI: 10.1088/0964-1726/6/2/007.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

[130] S. D. Snyder and N. Tanaka. “A Neural Network for Feedforward Controlled Smart Structures”. In: *IEEE* 4 (1993), pp. 373–378. ISSN: 1045-389X. DOI: 10.1177/1045389x9300400310.

[131] S. D. Snyder and N. Tanaka. “Active Control of Vibration Using a Neural Network”. In: *IEEE* 6 (1995), pp. 819–828. ISSN: 1045-9227. DOI: 10.1109/72.392246.

[132] T. T. Soong. “Introduction to Active Control”. en. In: ed. by T. T. Soong and M. C. Costantinou. Vienna: Springer, 1994, pp. 271–274. ISBN: 9783709130124. DOI: 10.1007/978-3-7091-3012-4_13. URL: https://doi.org/10.1007/978-3-7091-3012-4_13 (visited on 12/03/2024).

[133] A. Spadoni and M. Ruzzene. “Numerical and experimental analysis of the static compliance of chiral truss-core airfoils”. In: *Journal of Mechanics of Materials and Structures* 2.5 (July 2007), pp. 965–981. ISSN: 1559-3959. DOI: 10.2140/jomms.2007.2.965. URL: https://msp.org/jomms/2007/2-5/p09.xhtml (visited on 12/06/2024).

[134] M. G. Spencer, R. M. Sanner, and I. Chopra. “An Adaptive Neurocontroller for Vibration Suppression and Shape Control of a Flexible Beam”. In: 9 (1998), pp. 160–170. ISSN: 1045-389X. DOI: 10.1177/1045389x9800900302.

[135] G. E. Stavroulakis et al. “Design and robust optimal control of smart beams with application on vibrations suppression”. In: *Advances in Engineering Software*. Selected papers from Civil-Comp 2003 and AICivil-Comp 2003 36.11 (Nov. 2005), pp. 806–813. ISSN: 0965-9978. DOI: 10.1016/j.advengsoft.2005.03.024. URL: https://www.sciencedirect.com/science/article/pii/S0965997805001298 (visited on 07/23/2024).

[136] R. S. Sutton and A. G. Barto. *Reinforcement Learning:An Introduction*. 1992. DOI: 10.1007/978-1-4615-3618-5_1.

[137] M. Tanaka, S. Chonan, and Z. W. Jiang. “Force Control of a Flexible Finger with Distributed Sensors and Piezoelectric Actuators”. In: *Journal of Intelligent Material Systems and Structures* 7.3 (May 1996), pp. 301–306. DOI: 10.1177/1045389x9600700310.

[138] S. Thangudu. “Next Generation Nanomaterials: Smart Nanomaterials, Significance, and Biomedical Applications”. In: (2020), pp. 287–312. DOI: 10.1007/978-981-15-4802-4_15.

[139] H. F. Tiersten. *Linear Piezoelectric Plate Vibrations*. 1969, pp. 141–168. DOI: 10.1007/978-1-4899-5594-4_13.

[140] A. Turing. “COMPUTING MACHINERY AND INTELLIGENCE”. In: (1950). DOI: 10.1093/oso/9780198250791.003.0017.

[141] R. V. Ulijn. “Enzyme-responsive materials: a new class of smart biomaterials”. In: *Journal of Materials Chemistry* 16.23 (2006), p. 2217. ISSN: 1364-5501. DOI: 10.1039/b601776m.

[142] M. T. Valoor, K. Chandrashekhara, and S. Agarwal. “Self-adaptive vibration control of smart composite beams using recurrent neural architecture”. In: *International Journal of Solids and Structures* 38 (2001), pp. 7857–7874. ISSN: 0020-7683. DOI: 10.1016/s0020-7683(01)00125-1.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

[143] A. Vaswani et al. *Attention Is All You Need*. 2017. DOI: 10.48550/ARXIV.1706.03762.

[144] R. Vepa. “Chapter 4 - Biomimetic Robotics”. In: ed. by A. Lakhtakia and R. J. Martín-Palma. Boston: Elsevier, Jan. 2013, pp. 81–105. ISBN: 9780124159952. DOI: 10.1016/B978-0-12-415995-2.00004-0. URL: https://www.sciencedirect.com/science/article/pii/B9780124159952000040 (visited on 12/06/2024).

[145] W. Wakeland. “Bode compensator design”. In: *IEEE Transactions on Automatic Control* 21.5 (Oct. 1976), pp. 771–773. ISSN: 0018-9286. DOI: 10.1109/tac.1976.1101362.

[146] C. Wan and C. R. Bowen. “Multiscale-structuring of polyvinylidene fluoride for energy harvesting: the impact of molecular-, micro- and macro-structure”. In: *Journal of Materials Chemistry A* 5 (2017), pp. 3091–3128. ISSN: 2050-7488. DOI: 10.1039/c6ta09590a.

[147] Y.-K. Wan et al. “Control of structures using neural networks”. In: 4 (1995), A149–A157. ISSN: 0964-1726. DOI: 10.1088/0964-1726/4/1a/018.

[148] F.-Y. Wang. “The exact and unique solution for phase-lead and phase-lag compensation”. In: *IEEE Transactions on Education* 46.2 (May 2003), pp. 258–262. ISSN: 0018-9359. DOI: 10.1109/te.2002.808279.

[149] P. Wang. “Active vibration control in aspecific zone of smart structures”. PhD thesis. Ecole Centrale de Lyon, 2019.

[150] P. Wang et al. “Active vibration control in specific zones of smart structures”. In: *Control Engineering Practice* 84 (2019), pp. 305–322. ISSN: 0967-0661. DOI: https://doi.org/10.1016/j.conengprac.2018.12.005. URL: https://www.sciencedirect.com/science/article/pii/S0967066118301618.

[151] K. Yi et al. “Smart metacomposite-based systems for transient elastic wave energy harvesting”. In: *Smart Materials and Structures* 26.3 (Feb. 2017), p. 035040. ISSN: 0964-1726. DOI: https://doi.org/10.1088/1361-665X/aa5a5a.

[152] L. Zadeh. “Fuzzy sets”. In: *Information and Control* 8.3 (June 1965), pp. 338–353. DOI: 10.1016/s0019-9958(65)90241-x.

[153] Y. Zheng et al. “A piezo-metastructure with bistable circuit shunts for adaptive nonreciprocal wave transmission”. In: *Smart Materials and Structures* 28.4 (Mar. 2019), p. 045005. DOI: 10.1088/1361-665x/ab083c.

[154] J. G. Ziegler and N. B. Nichols. “Optimum Settings for Automatic Controllers”. In: *Transactions of the American Society of Mechanical Engineers* 64.8 (Dec. 2022), pp. 759–765. ISSN: 0097-6822. DOI: 10.1115/1.4019264.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Appendix A

1D Finite Element Model of a Smart Cantilever Beam: Numerical Datas

This appendix provides a numerical application of the 1 Dimensional Finite Element Model of the smart cantilever beam described in Chapter 2 with a matlab code available online in a repository : `https://github.com/mfebvre/PiezoBeam-OptDRL`

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Appendix B

Cantilever beam transfer function values for Chapter 4

This appendix provide poles and zeros values of the identified transfer functions from Chapter 4.

Table B.1: Pole and Zero values set for G_1

Pole	Zero
-62831.8530717961 + 0.00000000000000i	22333.7493581984 + 24241.9032933657i
-28410.7995685980 + 0.00000000000000i	22333.7493581984 - 24241.9032933657i
-102.324674940395 + 25671.9873959243i	-902.393677722916 + 32869.1197897130i
-102.324674940395 - 25671.9873959243i	-902.393677722916 - 32869.1197897130i
-14161.2805693114 + 12058.9901781808i	-382.770615104972 + 15571.2279654967i
-14161.2805693114 - 12058.9901781808i	-382.770615104972 - 15571.2279654967i
-423.996243560118 + 15449.9253538215i	-39.9113483038095 + 8411.01603931739i
-423.996243560118 - 15449.9253538215i	-39.9113483038095 - 8411.01603931739i
-53.7255612124484 + 7963.82135319923i	-16.0593916817720 + 2038.38069495669i
-53.7255612124484 - 7963.82135319923i	-16.0593916817720 - 2038.38069495669i
-12.1807219370647 + 1824.79739755134i	
-12.1807219370647 - 1824.79739755134i	

Table B.2: Pole and Zero values set for G_2

Pole	Zero
-53.0910194206690 + 7964.59589614937i	2746.42511952631 + 9045.91214678551i
-53.0910194206690 - 7964.59589614937i	2746.42511952631 - 9045.91214678551i
-2493.16089827939 + 5458.38220738164i	49.0400490814010 + 5033.47080791650i
-2493.16089827939 - 5458.38220738164i	49.0400490814010 - 5033.47080791650i
-3954.58648147516 + 0.00000000000000i	3971.22964431937 + 2585.89970012983i
-12.1578332789696 + 1824.85597977042i	3971.22964431937 - 2585.89970012983i
-12.1578332789696 - 1824.85597977042i	

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Appendix C

Measured and Identified Distributed System Transfer Function from Chapter 5

This appendix provides the transfer functions presented in Chapter 5 of the manuscript and shown in Figure C.1.

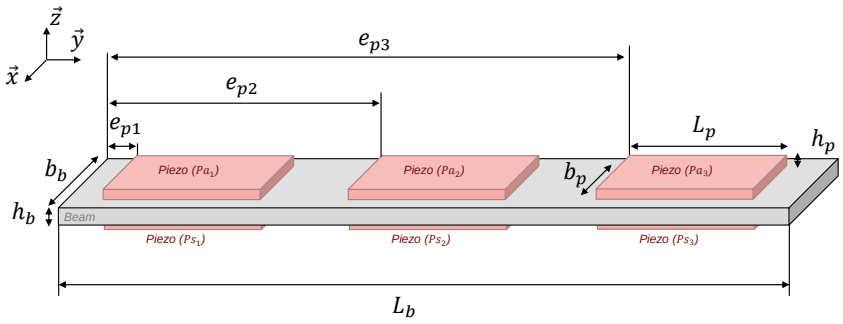
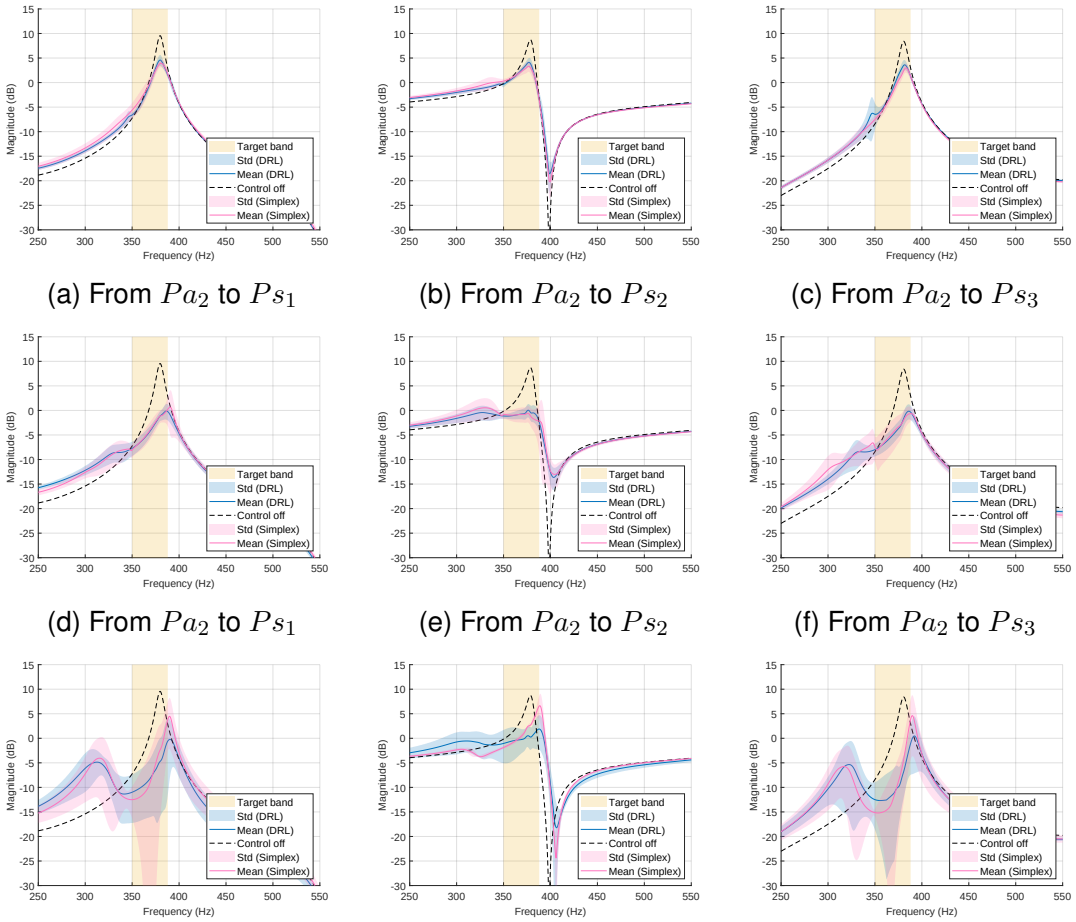


Figure C.1: Smart Beam diagram

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés



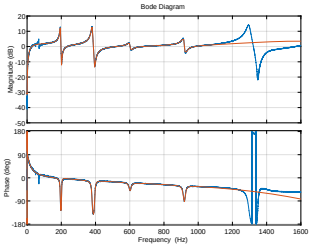
(a) From Pa_2 to Ps_1 (b) From Pa_2 to Ps_2 (c) From Pa_2 to Ps_3

(d) From Pa_2 to Ps_1 (e) From Pa_2 to Ps_2 (f) From Pa_2 to Ps_3

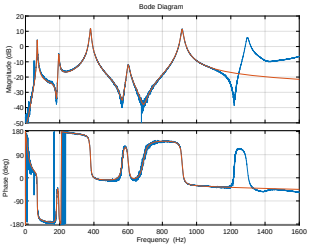
(g) From Pa_2 to Ps_1 (h) From Pa_2 to Ps_2 (i) From Pa_2 to Ps_3

Figure C.2: Bode plot for different controllers applied to the model: final reward between 9 and 13 (a-b-c), final reward between 13 and 17 (d-e-f) and final reward over 17 (g-h-i).

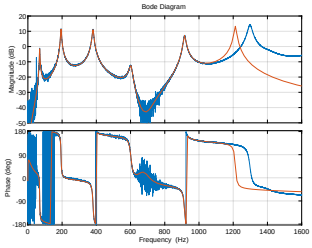
Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés



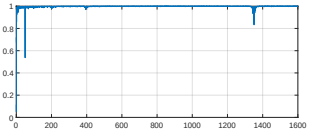
(a) $G_{1,1}$: From Pa_1 to Ps_1



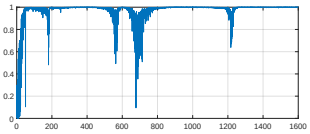
(b) $G_{2,1}$: From Pa_2 to Ps_1



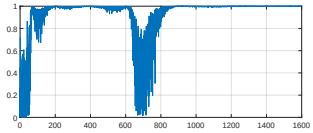
(c) $G_{3,1}$: From Pa_3 to Ps_1



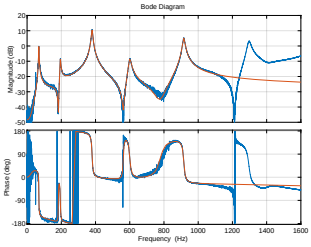
(d) Coherence: From Pa_1 to Ps_1



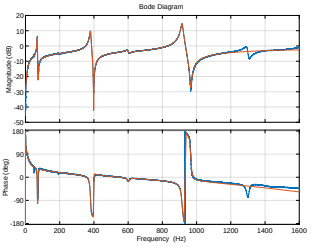
(e) Coherence: From Pa_2 to Ps_1



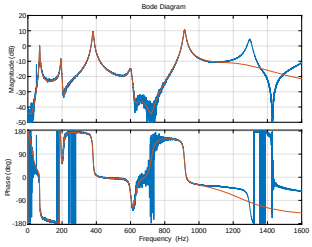
(f) Coherence: From Pa_3 to Ps_1



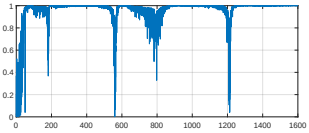
(g) $G_{1,2}$: From Pa_1 to Ps_2



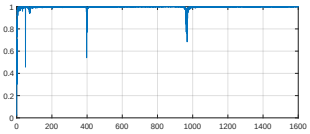
(h) $G_{2,2}$: From Pa_2 to Ps_2



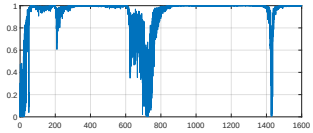
(i) $G_{3,2}$: From Pa_3 to Ps_2



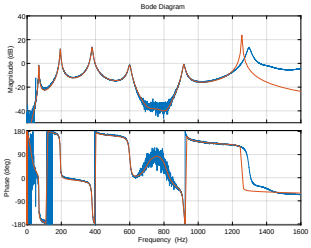
(j) Coherence: From Pa_1 to Ps_2



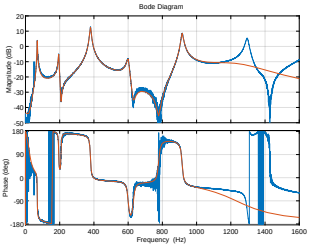
(k) Coherence: From Pa_2 to Ps_2



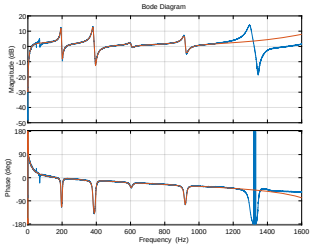
(l) Coherence: From Pa_3 to Ps_2



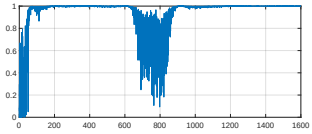
(m) $G_{1,3}$: From Pa_1 to Ps_3



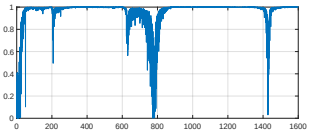
(n) $G_{2,3}$: From Pa_2 to Ps_3



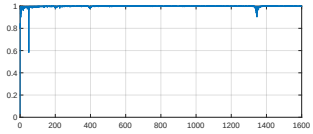
(o) $G_{3,3}$: From Pa_3 to Ps_3



(p) Coherence: From Pa_1 to Ps_3



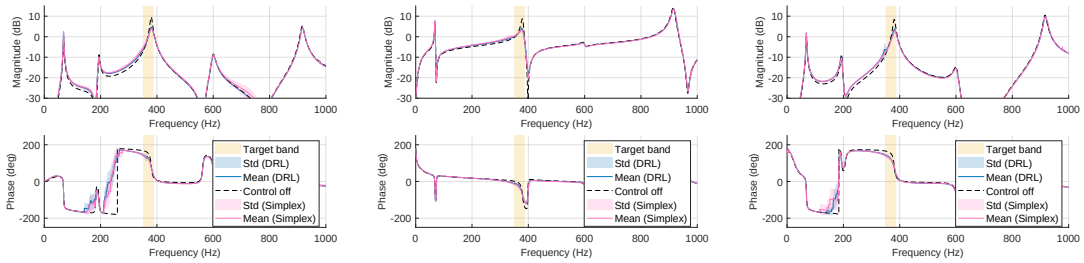
(q) Coherence: From Pa_2 to Ps_3



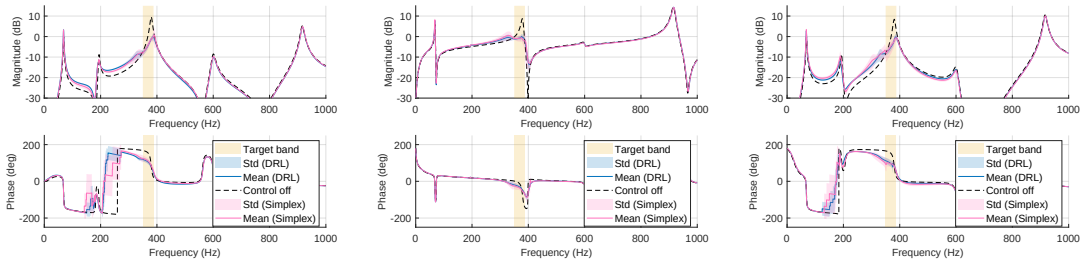
(r) Coherence: From Pa_3 to Ps_3

Figure C.3: Bode and coherence plot with transfer function estimated from measurements (blue line) and the identified model (red line) between 0-1500 $[Hz]$

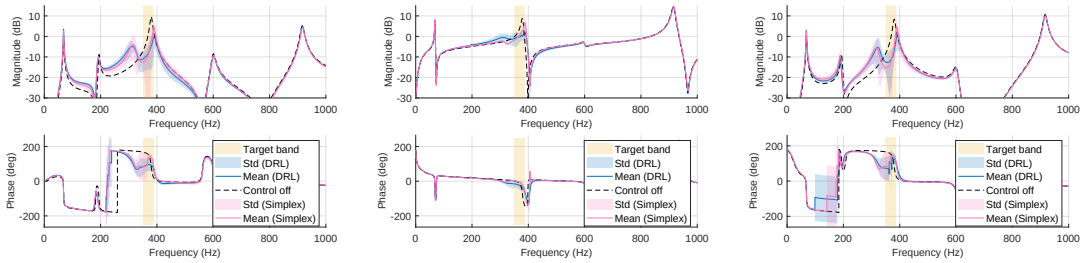
Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés



(a) From Pa_2 to Ps_1 (b) From Pa_2 to Ps_2 (c) From Pa_2 to Ps_3



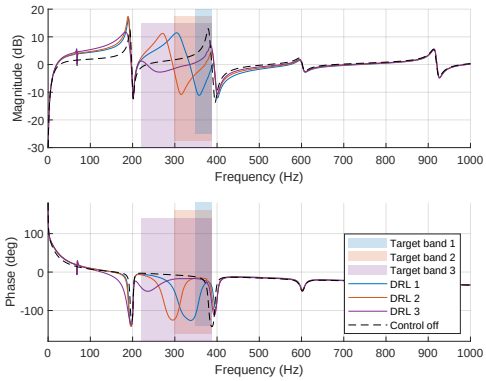
(d) From Pa_2 to Ps_1 (e) From Pa_2 to Ps_2 (f) From Pa_2 to Ps_3



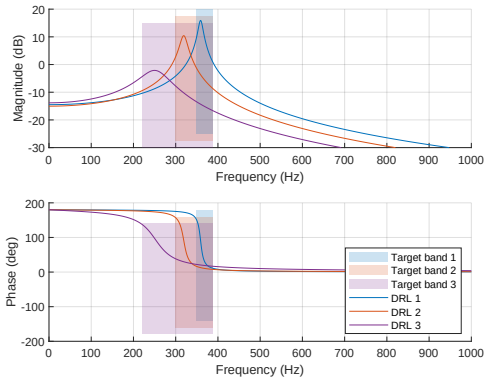
(g) From Pa_2 to Ps_1 (h) From Pa_2 to Ps_2 (i) From Pa_2 to Ps_3

Figure C.4: Bode plot for different controllers applied to the model: final reward between 9 and 13 (a-b-c), final reward between 13 and 17 (d-e-f) and final reward over 17 (g-h-i).

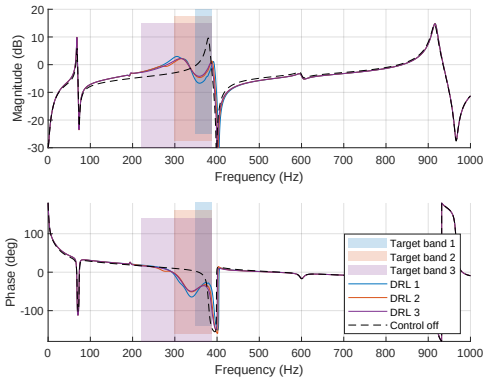
Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés



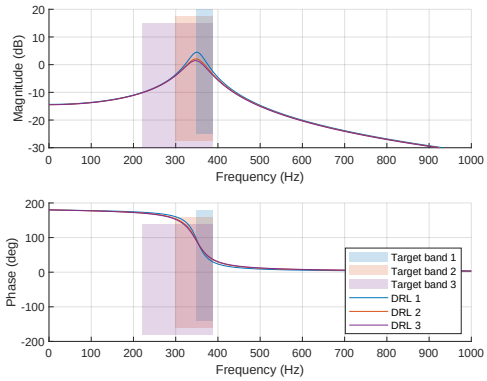
(a) From Pa1 to Ps1



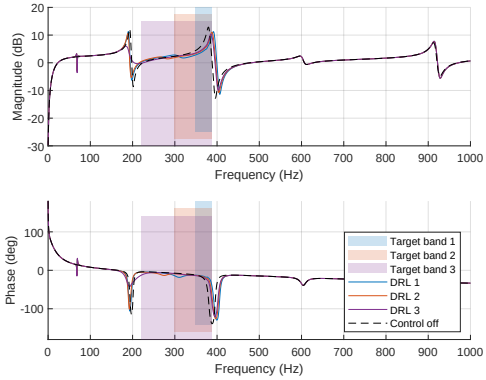
(b) Controller loop 1



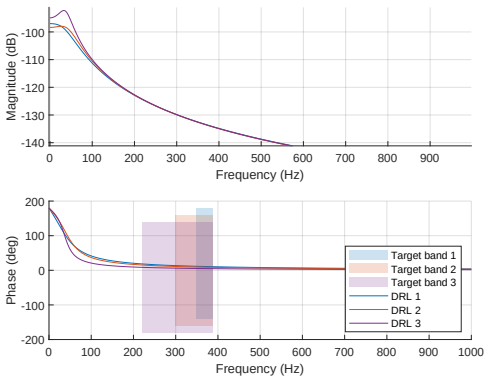
(c) From Pa2 to Ps2



(d) Controller loop 2



(e) From Pa3 to Ps3



(f) Controller loop 3

Figure C.5: Bode plot for three tuned controllers applied to the model with increasing target band frequency range.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

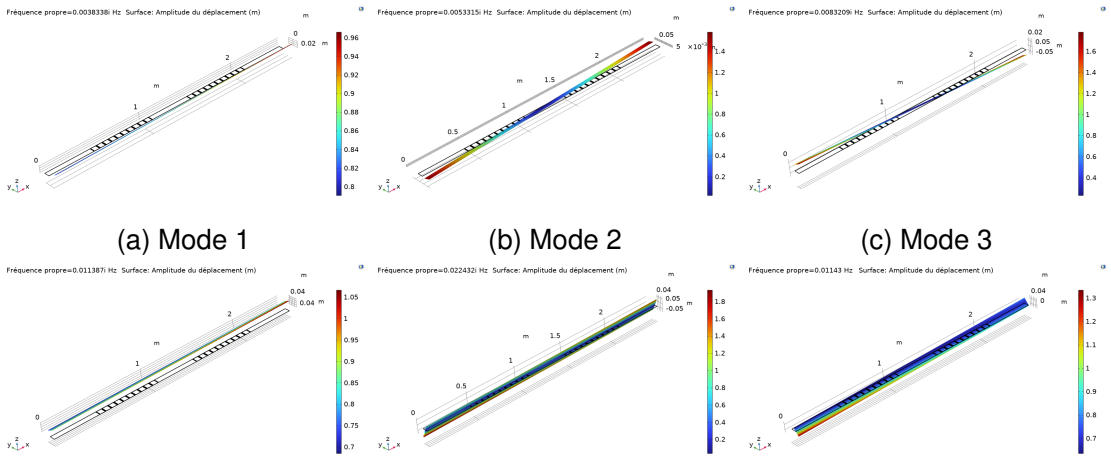
Appendix D

Modal distribution of the structure between 0 and 590 Hz: modal shapes

This appendix presents the modal analysis of a free end beam equipped with 20 pairs of transducers. The modal shapes between 0 and 590 Hz are classified as follows:

- Rigid modes: 6 in Figure D.1
- Bending modes in x-axis: 5 in Figure D.2
- Torsional modes in y-axis: 6 in Figure D.3
- Z-axis bending mode: 21 Figure in D.4

The modes identified during the measurements are only bending in the z-axis as shown in Figure 6.5.

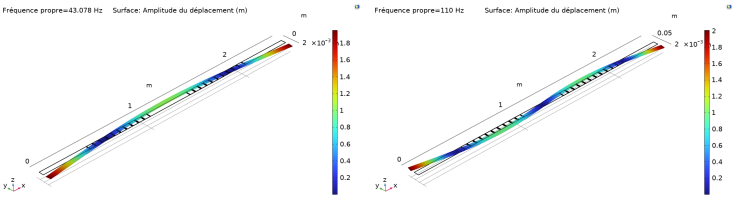


(a) Mode 1 (b) Mode 2 (c) Mode 3

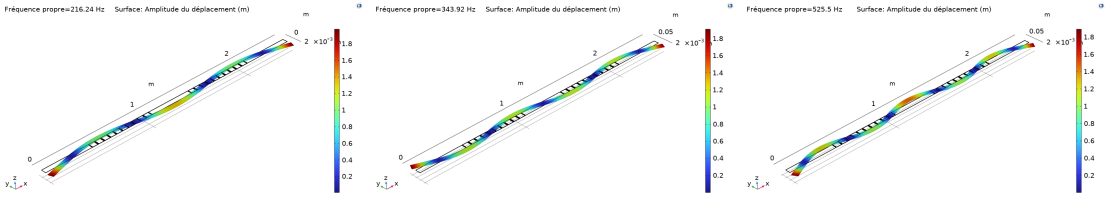
(d) Mode 4 (e) Mode 5 (f) Mode 6

Figure D.1: Shape of rigid body modes

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

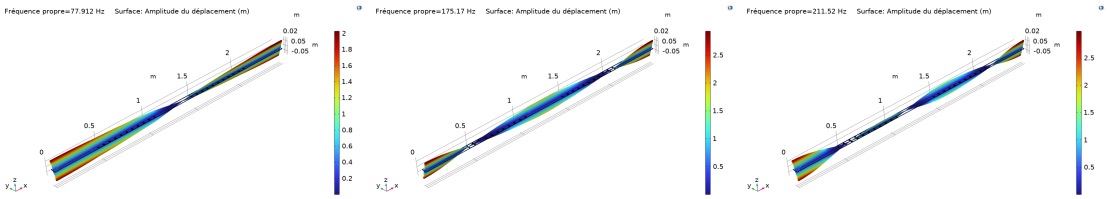


(a) Mode 12 (b) Mode 18

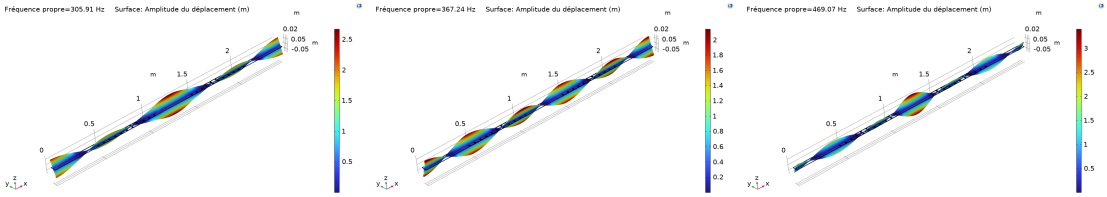


(c) Mode 24 (d) Mode 30 (e) Mode 36

Figure D.2: Bending modes in the x-axis



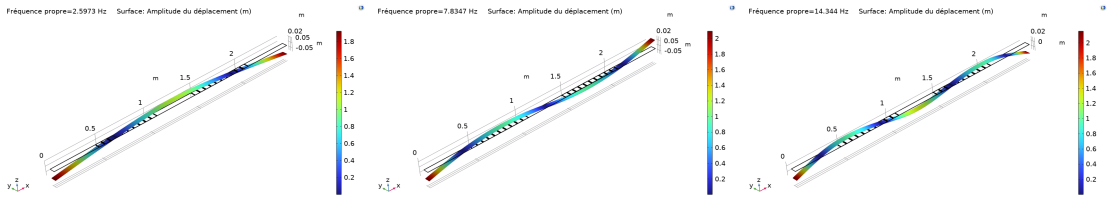
(a) Mode 15 (b) Mode 21 (c) Mode 23



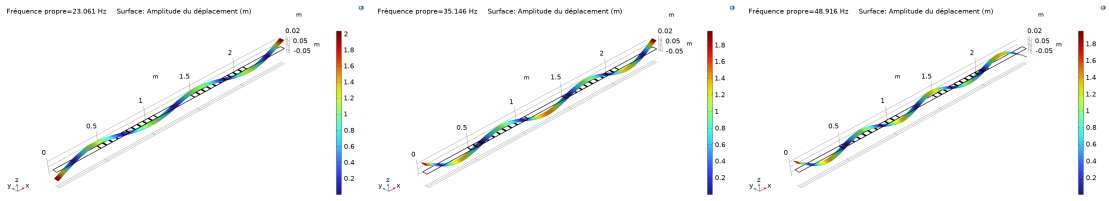
(d) Mode 28 (e) Mode 31 (f) Mode 34

Figure D.3: Torsion modes in the y-axis

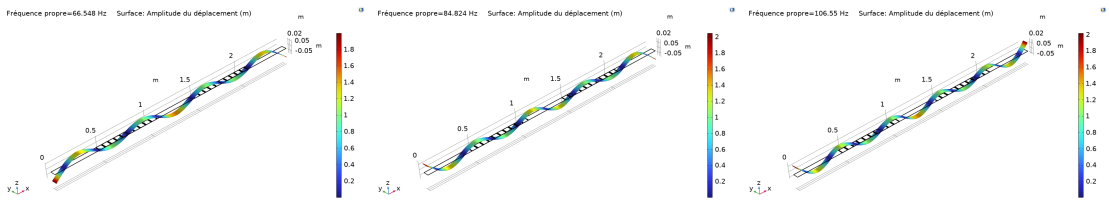
Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés



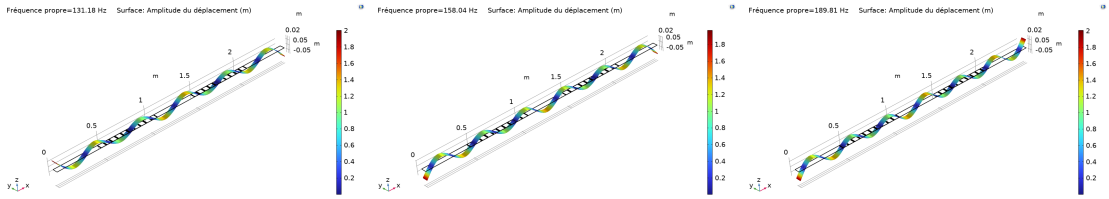
(a) Mode 7 (b) Mode 8 (c) Mode 9



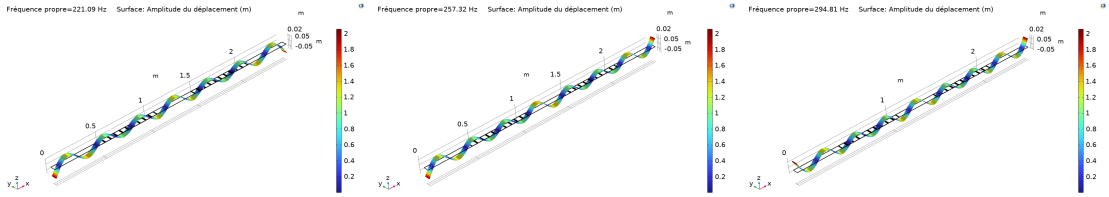
(d) Mode 10 (e) Mode 11 (f) Mode 13



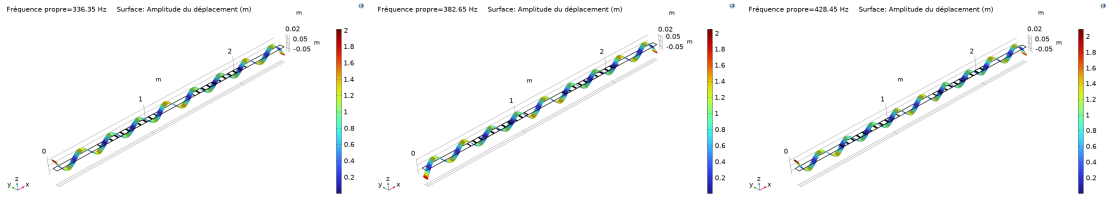
(g) Mode 14 (h) Mode 16 (i) Mode 17



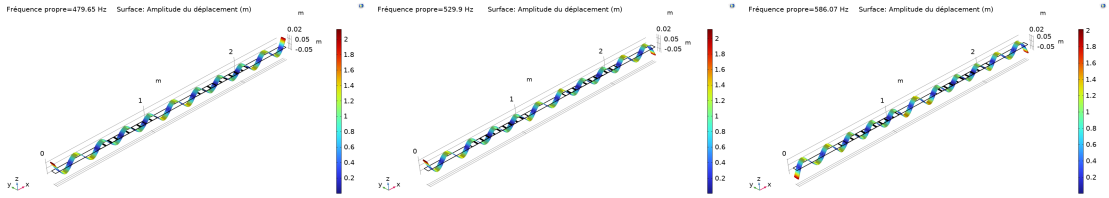
(j) Mode 19 (k) Mode 20 (l) Mode 22



(m) Mode 25 (n) Mode 26 (o) Mode 27



(p) Mode 29 (q) Mode 32 (r) Mode 33



(s) Mode 35 (t) Mode 37 (u) Mode 39

Figure D.4: Bending modes in the z-axis

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Appendix E

Shunt RC_{neg} Equivalent capacity demonstration

Mason or KLM equivalent circuit [125] for piezoelectric resonators provide representations as shown in Figure E.1. The example is run on transducer number 4 but is equivalent for all transducers.

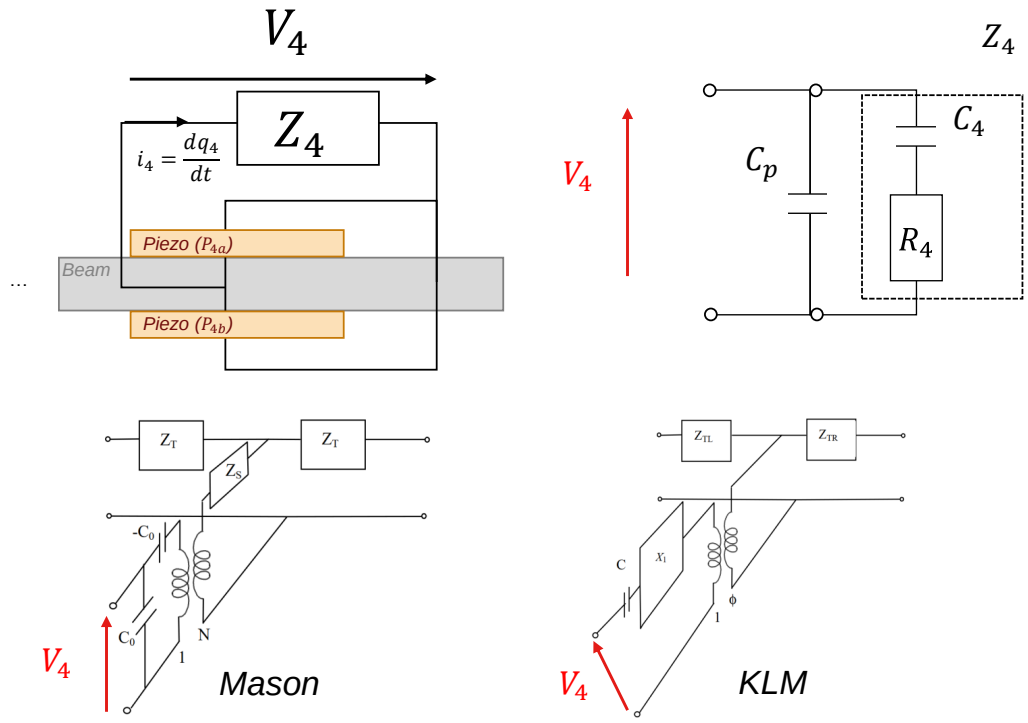


Figure E.1: Mason and KLM equivalent circuits for piezoelectric resonators in the thickness mode, schematics from [125], $Z_4 = R_4 + \frac{1}{C_4 s}$

For both cases, the shunt capacitance and the added negative capacitance can be assumed as in series:

- For the KLM representation, the demonstration is trivial.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

- For the Mason representation, the equations can be written using the voltage driver configuration formulation (Figure E.2):

$$Veq = I(\frac{1}{j\omega C_4} + \frac{1}{j\omega C_p}) \tag{E.1}$$

Where C_p is the transducer capacitance and C_4 is the shunt capacitance. The equivalent capacitance is written such as:

$$C_{eq} = \frac{1}{\frac{1}{C_4} + \frac{1}{C_p}} \tag{E.2}$$

From this formulation, V_4 is written such as:

$$V_4 = V_{eq}\frac{1}{1 + \frac{C_4}{C_p}} \tag{E.3}$$

This equations is represented in Figure E.3 with an identical behavior as the equivalent stiffness of the structure as a function of the capacity [21].

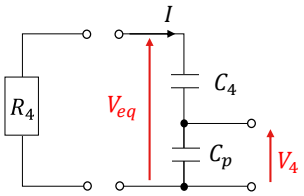


Figure E.2: Mason equivalent circuit in a voltage driver form with RC shunt.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

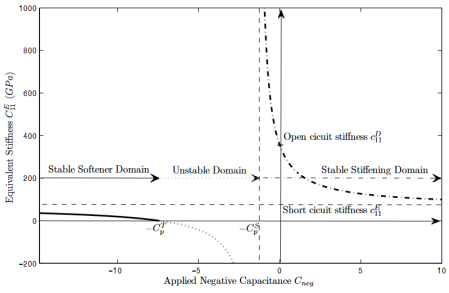
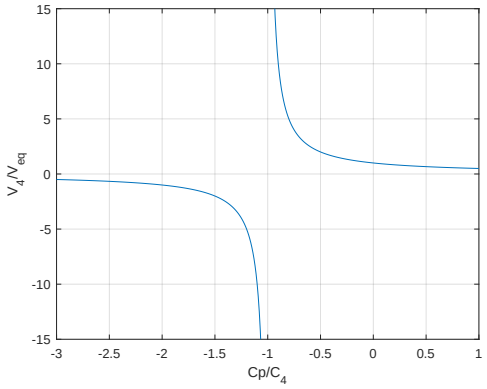


Fig. 3 The effective shunted piezoelectric composite stiffness [36] as a function of the connected negative capacitance term.

(a)



(b)

Figure E.3: Parallel between the effective shunted piezoelectric stifness from [21] (a) and the voltage driver behavior (b)

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Appendix F

Calculation of the $C0_2$-equivalent of the numerical tools used

This appendix provides a general idea of the environmental impact of this work by considering only the numerical systems used for three years. It takes into account the power consumption, the utilization rate and the manufacturing impact of the computing units. An optimistic approximation, considering $59.9gCO2e/kWh$, is calculated based on the work of INRIA [9] and with the help of Arnaud Duval, the Sustainable Development and Corporate Social Responsibility Committee of the LaMCoS laboratory. Since this experiment takes place in France, the consumption calculations are based on Enedis (`https://www.enedis.fr/`) information in aggrement with ADEME (`https://www.ademe.fr/`). The calculations for this thesis were done on a DELL laptop Dell (`https://www.dell.com/fr-fr`) and a computing servers (cluster), so these two parts are taken into account to estimate the C02 equivalent consumption in Table F.1.

Table F.1: $C0_2$ footprint of the digital part of this work

Equipment	Categpry	$C0_2$e	unit
Computing servers (cluster)	Manufacturing	1.91	g$C0_2$e for 1h.core
	Server electricity consumption	1.01	g$C0_2$e for 1h.core
	Room electrical consumption	1.01	g$C0_2$e for 1h.core
	Total	3.93	g$C0_2$e for 1h.core
DELL Laptop	Manufacturing (if used 6 years)	97	kg$C0_2$e
	Electricity consumption	33.8	kg$C0_2$e
TOTAL	Cluster	36.7	kg$C0_2$e for 9327h.core
	Laptop	130.8	kg$C0_2$e for 3 years
	Cluster + Laptop	167.5	kg$C0_2$e

It should be noted that some elements are not taken into account such as:

- Server Room Construction
- Construction, power consumption and space consumption associated with storage, surveillance or backup
- Emissions associated with staff managing the equipements (commuting, etc.)
- Network equipment

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

Readers have to keep in mind that taking into account everything is challenging with the actual tools and including new items will increase the carbon footprint estimation. This first estimation helps to compare the AI tool used in this work with chat GPT3 by using $C0_2$ equivalent consumption as reference. According to the Stanford Artificial Intelligence Index report [88] and David Patterson et al. [107], the initial training of chat GPT3 required 552 tons of $C0_2$ equivalent. The entire work consumed only 0.03% of the $C0_2$ equivalent used to train chat GPT3. Since the use of AI methodology is only at its beginning in the mechanical domain, it has the potential to improve to reach the efficiency of chat GPT in future research with appropriate resources.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

List of Figures

1.1 Aircraft wing design evolution with time through four examples 14
1.2 Example of adaptive structures applied to optical technologies. 14
1.3 Piezoelectric smart structures literature overview 15
1.4 Diagram of the experimental set-up used by the Curie brothers to demonstrate the indirect piezoelectric effect [24] (p.40): pressure measurement system (a, b ,c), system producing the phenomenon (a', b', c'), Holtz machine (M), ball micrometer poles (B), ground (T and T'), battery (P) of the eletrometer (e). 16
1.5 Classification of smart materials based on the physics of external stimuli . 17

2.1 Piezoelectric material with no strain (a), shrinkage under stress with direct effect (b), and strain under applied voltage with converse effect (c). 24
2.2 A body moving in reference space O, where r is the distance from the origin of reference O, ρ is the density of the body, $\dot{u}$ is the velocity vector, f is the body force vector, n is the normal vector, and $t(n)$ is the contact force vector. 26
2.3 Deformation of an arbitrary volume element with a representation of the displacement vectors between an initial state at time t_0 and a current state at time t_1 28
2.4 Architecture and orthotropic axes of a planar piezoelectric layer with electrodes being perpendicular to the polarization axis (the substrate is not represented). 32
2.5 Active Control Methods Classification 34
2.6 Feedback control law $C(s)$ applied on system $G(s)$ 35
2.7 Visualization of the stability of the system $G_{FB}(s)$ where $C(s) = 40$ using: Complex Plane (a), Nyquist (b), Bode (c) and Nichols (d) plots. 37
2.8 Closed-loop Bode (a), open-loop Bode (b), impulse (c), and step (d) responses of the $G(s)$ system with proportional feedback controllers for different gains. 38
2.9 Closed-loop Bode (a), open-loop Bode (b), impulse (c), and step (d) responses of the $G(s)$ system with proportional integrator feedback controllers for different gains. 39
2.10 Closed-loop Bode (a), open-loop Bode (b), impulse (c), and step (d) responses of the $G(s)$ system with Derivative (D) feedback controllers for different gains. 40
2.11 Root Locus of the system G with a Derivative controller and red (x) poles of the controlled system with $K_d = 21$. 41
2.12 Bode (a) and Pole Zero (b) plots of Lead or Lag compensator for various gains. 42

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

2.13 Closed-loop Bode (a), open-loop Bode (b), impulse (c), and step (d) responses of the $G(s)$ system with lead or lag compensators for different gains. 43
2.14 Bode plots of PPF controllers for different parameters 44
2.15 Root-locus plots of PPF controllers for different parameters: PPF tuned to target the first mode (a), and PPF tuned to target the second mode (b). 45
2.16 Closed-loop Bode (a), open-loop Bode (b), impulse (c), and step (d) responses of the $G(s)$ system with lead or lag feedback controllers for different gains. 45
2.17 Closed-loop bode (a), with a zoom in frequency between 2000 and 4000 $rad.s^{-1}$(b) responses of the $G(s)$ system with the tuned studied feedback controllers. 46
2.18 Bode response of the system $G(s)$ with (blue) short-circuit and (red) open-circuit piezoelectric electrodes zoom around the first mode. 48
2.19 Bode of the system G(s) first mode (a) for a resistive shunt varying between 0 (short circuit) and infinity (open circuit) and (b) the corresponding root locus. 48
2.20 Hagood Transfer functions for a 1-DOF [43] compared to the system G(s) first mode for a resistive inductive shunt at various tuning (b) normalized with the short-circuit frequency and (c) not normalized. 50
2.21 Bode (a) and Complex plane (b) of system G(s) response with various optimal shunt 51
2.22 From IBM website: Illustration of gradient descent algorithm with various case: (a) small and large learning rate (b) local minimum and Saddle-point. 52
2.23 From Cheng [16]: Illustration of simplex geometric variation with cases . . 55
2.24 From Matanga [89]: Illustration of generation evolution. 56
2.25 Connectionnist versus Symbolic Approaches Evolution between 1935 and 2018 by Cardon et al. [13] 58
2.26 Organization of a perceptron: (a) Brain cell (b) Rosenblatt [118] and (c) modern representation 59
2.27 From FIDLE: Perceptron used as a classifier on (a) iris plants dataset from Fisher [29], with (dark color) train set and (light color) prediction set. . . . 60
2.28 Neural Network schematic representation 60
2.29 Artificial intelligence domains mapping 63
2.30 Reinforcement Learning algorithm: scheme 64

3.1 The reference system: Euler-Bernoulli cantilever beam with two collocated piezoelectric transducers. 72
3.2 System $G(s)$ slightly damped: Bode diagram (a) and pole (blue $\times$) / zero(red o) pattern (b) of the structure with collocated actuator and sensor with $N_b = 10$. 74
3.3 Control loop block diagram 75
3.4 Reinforcement Learning Algorithm applied on the 1D smart beam 76
3.5 Reward: (blue line) variation with resistance and (black "+") resistance tuning using Hagood (Chapter 2 Part 2.1.2.2). 78
3.6 Training: (a) TRPO, (b) PPO. 80
3.7 Control Performances on the example system: (a) Bode response of G(s), (b) Complex plan, (c) Impulse and (d) Noise response. 81

4.1 Scheme: cantilever beam dimensions 84

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

4.2 Finite element model of the cantilever beam with different meshes: (a) extremely coarse, (b) coarser, (c) normal, (d) finer. 86
4.3 Influence of the mesh on the short and open circuit frequencies (left axis) and computation time (right axis). 86
4.4 Coupling coefficient variations with beam dimensions: (a) length, (b) width, (c) thickness ratio where $e_p = 0$ and (d) transducers position. 87
4.5 Mode shape visualization: (a) mode 1 (first bending mode), (b) mode 2 (second bending mode), (c) mode 3 (first torsion mode). 88
4.6 Key steps for sticking the transducer to the beam: (a) surface sanding , (b) alcohol cleaning, (c) applying and spreading the adhesive and, (d) vacuum drying. 89
4.7 Experimental setup: beam 90
4.8 Experimental setup: data acquisition systems 90
4.9 Bode diagram with transfer function estimated from measurements (blue line) and the identified model (red line) between 0-5000 $[Hz]$: (4.9a) G_1 and (4.9b) G_2 with (4.9c-4.9d) their corresponding spectral coherence . . . 93
4.10 Time measurements from the experimental setup (blue line) and identified model (red dashed line) during 25 sec and between seconds 4.5 and 4.6: piezoelectric (4.10a-4.10c) and laser (4.10b-4.10d) sensors 94
4.11 Control loop block diagram 95
4.12 Root locus from identified controlled model $G_1C(s)$ open loop with (a) $z_0 = 0$ and $p_0 = 6280$ and (b) a zoom around the two first poles and zeros (o) poles ($\times$) estimated from measurement on the experimental setup when control is on for different values of K_c from $K_{c_{min}} = -0.04$ ($\times$ dark red) to $K_{c_{max}} = 0.5$ ($\times$ dark blue) 97
4.13 Average reward values variations with derivative K_d control values (red line) and the 0.95 confidence interval (blue area) for 10 executions. 99
4.14 Reward R_1 (green line) and R_2 (orange line) variations with K_c for different lead filter z_0 and p_0 combinations : $z_0 = 6280$, $p_0 = -6280$ ($\times$); $z_0 = -31400$, $p_0 = -6280$ ($+$); $z_0 = 31400$, $p_0 = -12560$ ($\triangle$) and $z_0 = -6280$, $p_0 = -21980$ ($*$) tested on the model 100
4.15 Training influence with hyperparameter variations: (a) entropy Loss weight and (b) learning rate 101
4.16 Reward R_1 after 600 episodes, mean and standard deviations over 10 different NN initialization. 103
4.17 Average reward value variations with episode number (dark blue line), the 0.95 confidence interval (blue area) to tune a derivative controller (10 Trainings). 104
4.18 Transfer functions (a) $G_1(s)$ and (b) $G_2(s)$ with control off $Kd_{None} = 0$ (blue), control on with $Kd_{TRPO_{best}} = 7 \times 10^{-5}$ (red) and control on with $Kd_{TRPO_{avg}} = 5.29 \times 10^{-5}$ (orange) computed from x_1 and x_2 experimental measurements. 106
4.19 Figure 4.18b zoom between 200 and 400 Hz. 106
4.20 Training on model with disturbance v_{d1}: mean reward over all training (dark blue), the corresponding standard deviation area at each episode step (light blue), the equivalent reward when the control is off (green - -) and the equivalent reward with H_2 norm minimization controller parameters (red .-). 110

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

4.21 Training on model with chirp disturbance v_{d2}: mean reward over all training (dark blue), the corresponding standard deviation area at each episode step (light blue), the equivalent reward when the control is off (green - -) and the equivalent reward with H_∞ norm minimisation controller parameters (red .-). 110
4.22 Case v_{d1}: system passive response (green - -), system controlled response with parametric H_2 norm minimisation (red .-), mean (dark blue) and standard deviation (light blue) over all system controlled response with DRL tuned controller: (a) model and (c) experiment $G_{c,1}$ bode diagram and (b) model and (d) experiment $G_{c,2}$ bode diagram. 113
4.23 Case v_{d2}: system passive response (green - -), system controlled response with parametric H_∞ norm minimization (red .-), mean (dark blue) and standard deviation (light blue) over all system controlled response with DRL tuned controller: (a) model and (c) experiment $G_{c,1}$ bode diagram and (b) model and (d) experiment $G_{c,2}$ bode diagram. 115
4.24 Flowchart: restarting or continuing the training to adapt from input noise v_{d1} (Reward 1) to v_{d2} (Reward 2) 118
4.25 Training on model with noise disturbance v_{d1} (a) with and without pre-training with chirp disturbance v_{d2} and (b) a zoom on the 700 first episodes of the training on case v_{d1}. 119
4.26 Flowchart: multi epoch training to adapt from input noise v_{d1} (Reward 1) and v_{d2} (Reward 2) 120
4.27 Training on model with noise v_{d12} and chirp disturbance v_{d2}: mean reward over all training (dark blue), the corresponding standard deviation area at each episode step (light blue), the equivalent reward when the control is off (green - -) and the equivalent sum of rewards $R_1 + R_2$ with H_2 and H_∞ norm minimization controller parameters (red .-). 120

5.1 Beam dimensions diagram 125
5.2 Mode shape visualization: (a) Mode 7 at 72 Hz (first bending mode), (b) Mode 8 at 194 Hz (second bending mode), (c) Mode 9 at 299 Hz (first torsion mode), (d) Mode 10 at 380 Hz (third bending mode), (e) Mode 11 at 600 Hz (second torsion mode), (f) Mode 12 at 603 Hz (fourth bending mode), (g) Mode 13 at 903 Hz (third torsion mode), (h) Mode 14 at 915 Hz (fifth bending mode). 127
5.3 Modal coupling coefficient variations with colocalized transducer positions e_{p1} and e_{p3}: (a) mode 7, (b) mode 8, (c) mode 10, and (d) mode 10 with non-colocalized transducers where $e_{p1} = 0$ and $e_{p3} = 0.34$. 128
5.4 Experimental setup: beam with piezoelectric transducers (top view). 129
5.5 Experimental setup: data acquisition systems 129
5.6 Bode and coherence plot with transfert function estimated from measurements (blue line) and the identified model (red line) between 0-1500 $[Hz]$. 131
5.7 Time measurements from the experimental setup (blue line) and identified model (red dashed line) during 30 sec (a-b-c) and between seconds 6.0 and 6.02 (c-d-e). 132
5.8 Control loop block diagram 133

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

5.9 Bode plot for different controllers (Table 5.5) applied to the system: comparison of model (- -) and experimental (–) measurements where $\xi = 0.23$, $\omega = 2.39E03$, $K_c = 0.05E6$ (blue) , $K_c = 0.1E6$ (red), $K_c = 0.3E6$ (orange), $K_c = 0.5E6$ (green), $K_c = 0.8E6$ (cyan) and $K_c = 2E6$ (magenta). . 135
5.10 Reward R_1 variation with control law parameters set as identical for all control loops: $\omega_{1/2/3} = 380 \times 2\pi$ (a), $\xi_{1/2/3} = 0.55$ (b), $Kc_{1/2/3} = 1e6$ (c-d). 138
5.11 Flowcharts 140
5.12 Flowchart: DRL and Simplex with a single initialization 141
5.13 Proportion of tested cases according to the reward obtained at the end of the calculation (considering 85 interdependent cases tested) 141
5.14 Means and standard deviations of the number of environmental evaluation shared according to the reward obtained at the end of the calculation (considering 85 interdependent tested cases). 142
5.15 DRL reward variation with episodes grouped by final reward value achieved: between 9 and 13 (a), between 13 and 17 (b), and over 17 (c) (The figures do not represent simplex variations with an episode; it only represents the terminal state.) 143
5.16 Bode plot for different controllers applied to the model: final reward between 9 and 13 (a-b-c), final reward between 13 and 17 (d-e-f) and final reward over 17 (g-h-i). 144
5.17 Flowchart: continuing the training strategy 145
5.18 Reward variation with episodes using the continuing training strategy. (The manual settings are defined in part 5.1.3 146
5.19 Bode plot for three tuned controllers applied to the model with increasing target band frequency range. 147
5.20 Reward R_4 variation with control law parameters set as identical for all control loops: $\omega_{1/2/3} = 380 \times 2\pi$ (a), $\xi_{1/2/3} = 0.55$ (b), $Kc_{1/2/3} = 1e6$ (c). . 149
5.21 Flowchart: restarting or continuing the training strategy 150
5.22 Reward variation with episodes (a-c-e) and Bode plots (b-d-f) for the 20 best tuned controllers with DRL and simplex methods applied to the model for different strategies and reward definitions. 151
5.23 Bode plots for 10 tuned controllers with DRL measured on the experimental setup. 152

6.1 Experimental setup: beam with piezoelectric transducers. 157
6.2 Experimental setup: data acquisition systems 157
6.3 Bode from the force F_0 at the shaker location to the displacement measured at transducer P_1 location. 158
6.4 Beam dimensions diagram 159
6.5 Mode shape visualization: first bending mode (a), fifth bending mode (b) ninth bending mode (c). 163
6.6 Shunt Scheme 166
6.7 Control loop block diagram 166
6.8 Reward variations with Resistive Shunt 169
6.9 Reward variations with $R - C_{neg}$ Shunt 170
6.10 Bode R same Analytic tuning 171
6.11 Bode RC_{neg} same Analytic tuning 172
6.12 Genetic Algorithm diagram 173

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

6.13 DRL Training Process . . . 175
6.14 Training Identical case with DRL0 (a), Distinct case with DRL1 and DRL3 (b): R shunt . . . 175
6.15 Bode from F_{t1} to w_1 (a), with zoom on the first mode (b), the fifth (c) and the ninth (d) bending mode: R shunt - Identical case . . . 176
6.16 Bode from Ft_1 to w_1 (a), with zoom on the first mode (b), the fifth (c) and the ninth (d) bending mode: R shunt - Distinct case . . . 177
6.17 Training DRL0 (a), DRL1 and DRL3 (b), DRL1 with a zoom between episodes 1180 and 1270 (c) DRL2 (d): RC_{neg} . . . 179
6.18 Bode from Ft_1 to w_1 (a), with zoom on the first mode (b), the fifth (c) and the ninth (d) bending mode: RC_{neg} shunt - Identical case . . . 180
6.19 Bode from Ft_1 to w_1 (a), with zoom on the first mode (b), the fifth (c) and the ninth (d) bending mode: RC_{neg} shunt - Distinct case . . . 181

C.1 Smart Beam diagram . . . 211
C.2 Bode plot for different controllers applied to the model: final reward between 9 and 13 (a-b-c), final reward between 13 and 17 (d-e-f) and final reward over 17 (g-h-i). . . . 212
C.3 Bode and coherence plot with transfer function estimated from measurements (blue line) and the identified model (red line) between 0-1500 $[Hz]$. 213
C.4 Bode plot for different controllers applied to the model: final reward between 9 and 13 (a-b-c), final reward between 13 and 17 (d-e-f) and final reward over 17 (g-h-i). . . . 214
C.5 Bode plot for three tuned controllers applied to the model with increasing target band frequency range. . . . 215

D.1 Shape of rigid body modes . . . 217
D.2 Bending modes in the x-axis . . . 218
D.3 Torsion modes in the y-axis . . . 218
D.4 Bending modes in the z-axis . . . 219

E.1 Mason and KLM equivalent circuits for piezoelectric resonators in the thickness mode, schematics from [125], $Z_4 = R_4 + \frac{1}{C_4 s}$ . . . 221
E.2 Mason equivalent circuit in a voltage driver form with RC shunt. . . . 222
E.3 Parallel between the effective shunted piezoelectric stifness from [21] (a) and the voltage driver behavior (b) . . . 223

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

List of Tables

2.1 Constitutive Piezoelectric Equations from the IEEE Standard [92] 25
2.2 Piezoelectric equation variables definition 25
2.3 Gain and Phase margins of the open loop system $G(s)$ with Proportional controllers. 38
2.4 Gain and Phase margins of the open loop system $G(s)$ with Proportional Integrator controllers. 39
2.5 Gain and Phase margins of the open loop system $G(s)$ with Derivative (D) controllers. 41
2.6 Gain and Phase margins of the open loop system $G(s)$ with Lead or Lag compensators. 43
2.7 Gain and Phase margins of the open-loop system $G(s)$ with Positive Position Feedback controllers. 44

3.1 Constant Definitions 73
3.2 Elementary matrix Definition 73
3.3 Hyperparameter definition for PPO and TRPO algorithms 79

4.1 Constant definitions 85
4.2 Setup dimensions 88
4.3 Experimental devices 91
4.4 Coupling coefficients on the first mode 91
4.5 TRPO algorithm characteristics 101
4.6 Actor and critic neural network definition 103
4.7 Training results: derivative controller tuning 105
4.8 Derivative control law values and performances. 105
4.9 Training results after 20000 episodes and control law values and performances on model: v_{d1} case. 108
4.10 Training results after 30000 episodes and control law values and performances on model: v_{d2} case. 109
4.11 Case 1: lead control Law values and performances measured on the experimental setup 112
4.12 Case 2: lead control law values and performances measured on the experimental setup 114

5.1 Definition of the constants 126
5.2 Setup dimensions 127
5.3 Experimental devices 129
5.4 Model error quantification 132
5.5 Model error quantification with for PPF feedback control ($\xi = 0.23$ and $\omega = 2.39E03$) 134

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

5.6 Control Law values and reward computed from measurements on the experimental setup (Selected examples DRL 1, 3, 5 and 10 out of 10 training sessions) . . . 153

6.1 Modal parameters estimated from measurements from Figure 6.3 transfer function . . . 159
6.2 Transcuders capacitance: measurements . . . 160
6.3 Definition of the constants . . . 161
6.4 Modal parameters estimated from the model and compared with measurements from Table **??** . . . 165
6.5 Actor and critic neural network definition . . . 174
6.6 Shunt values and reward computed from the numerical model: R shunt . . . 178
6.7 Shunt values and reward computed from the numerical model: RC_{neg} . . . 182

B.1 Pole and Zero values set for G_1 . . . 209
B.2 Pole and Zero values set for G_2 . . . 209

F.1 $C0_2$ footprint of the digital part of this work . . . 225

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés

FOLIO ADMINISTRATIF

THESE DE L'INSA LYON, MEMBRE DE L'UNIVERSITE DE LYON

NOM : **FEBVRE** DATE de SOUTENANCE : **17/02/2025**

Prénoms : **Maryne**

TITRE : **L'intelligence artificielle pour l'optimisation du contrôle distribué : application aux réseaux de transducteurs des structures intelligentes.**

NATURE : **Doctorat** Numéro d'ordre : **2025ISAL0014**

École Doctorale : **MÉCANIQUE, ÉNERGÉTIQUE, GÉNIE CIVIL, ACOUSTIQUE**

Spécialité : **MÉCANIQUE**

RÉSUMÉ :
Les matériaux intelligents, comme les transducteurs piézoélectriques, sont devenus essentiels en ingénierie moderne pour des applications telles que le contrôle des vibrations, la récupération d'énergie et la propagation des ondes. Ces éléments multiphysiques permettent de développer des structures intelligentes adaptatives, capables d'interagir avec leur environnement, et de résoudre des problématiques liées à l'instabilité et à la fatigue des matériaux. Cependant, l'optimisation de ces systèmes devient de plus en plus complexe à mesure que le nombre de transducteurs et de paramètres ajustables augmente.
Cette thèse explore l'optimisation du contrôle des vibrations dans les structures intelligentes à l'aide de l'apprentissage par renforcement profond (DRL pour Deep Reinforcement Learning). Plusieurs lois de contrôle actif ou passif sont appliquées aux transducteurs piézoélectriques. Le réglage de ces lois par DRL est comparé à des méthodes d'optimisation traditionnelles telles que le simplex et les algorithmes génétiques. L'efficacité est évaluée en termes d'atténuation des vibrations, de stabilité structurelle et de performance de calcul. Des analyses modales, à la fois numériques et expérimentales, sont effectuées pour valider la faisabilité du contrôle sur diverses structures, allant de modèles unidimensionnels basés sur des éléments finis à des réseaux complexes de transducteurs.
Les résultats mettent en évidence l'efficacité du DRL pour ajuster des lois de contrôle en boucle fermée multi-paramètres tout en tenant compte de fonction d'optimisation non linéaires incluant des contraintes de stabilité. Cependant, des défis tels que l'aléa dans l'entraînement et la divergence sont surmontés grâce à des stratégies basées sur la mémoire, renforçant la robustesse et l'adaptabilité aux variations environnementales. Ce travail fait progresser les méthodes basées sur l'intelligence artificielle pour le contrôle des structures intelligentes distribuées, établissant un lien entre les domaines de l'intelligence artificielle et des matériaux adaptatifs.

MOTS-CLÉS : **Smart structure, réseau de transducteur piézoélectrique, Intelligence Artificielle.**

Laboratoire(s) de recherche : **Laboratoire de Mécanique des Contacts et des Structures - LaMCoS**
Laboratoire de Tribologie et Dynamique des Systèmes - LTDS

Directeur de thèse : **Simon Chesné**
Co-Directeur de thèse : **Manuel Collet**

Président du Jury : **Yann Meyer**

Composition du Jury :
Berardengo Marta, Dion Jean-Luc, Gardonio Paolo, Meyer Yann, Chesné Simon, Collet Manuel

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2025ISAL0014/these.pdf © [M. Febvre], [2025], INSA Lyon, tous droits réservés