



EURODINAME III - An International Symposium on Dynamic
Problems of Mechanics
June 8-11, 2026 - Giens Peninsula, FR

EURODINAME-2026-13458

REINFORCEMENT LEARNING FOR REACHING STABLE CONTROL OF ELECTROACOUSTIC ABSORBERS

Arthur Diniz Flor Torquato Fernandes

University of Naples Federico II, 80125, Naples, Italy

Université Marie et Louis Pasteur, SUPMICROTECH, CNRS, institut FEMTO-ST, F- 25000 Besançon, France

arthur.fernandes@femto-st.fr

Leonardo De Paula Silva Ferreira

Rafael de Oliveira Teloli

Morvan Ouisse

Université Marie et Louis Pasteur, SUPMICROTECH, CNRS, institut FEMTO-ST, F- 25000 Besançon, France

leonardo.ferreira@femto-st.fr, rafael.teloli@femto-st.fr, morvan.ouisse@femto-st.fr

Emanuele De Bono

Sergio De Rosa

Giuseppe Petrone

University of Naples Federico II, 80125, Naples, Italy

emanuele.debono@unina.it, derosa@unina.it, giuseppe.petrone@unina.it

Abstract. In recent years, active noise control approaches have gained growing interest in the market, being used in devices such as earbuds and headphones. One promising direction involves Electroacoustic Absorbers (EAs), systems that combine a loudspeaker diaphragm with collocated microphones. By monitoring the pressure inside a cavity and applying a control algorithm, the absorber adjusts the velocity at the resonator's surface to enforce a target acoustic impedance. This enables custom absorptions at specific frequency targets.

This type of control differs from conventional active noise control by not employing the concept of anti-noise. Instead of creating a quiet zone, these devices modify the boundary conditions of a chamber, much like passive solutions such as porous materials or Helmholtz resonators.

Although a specific impedance can currently be targeted, these systems remain incapable of self-adapting their control algorithms to different conditions and are prone to instability depending on the target impedance. To overcome these limitations, this work explores the use of reinforcement learning (RL) to improve adaptability and robustness under changing environments.

Keywords: Reinforcement learning, Acoustic impedance control, Harmonic absorption, Electroacoustic absorption

1. INTRODUCTION

Acoustic noise is a persistent problem that affects fields as varied as refrigeration systems, music production facilities, and wind energy. Noise is also a source of regulatory pressure. To mitigate this problem, noise-absorption techniques are usually employed.

Among the current noise-absorption solutions, there are passive approaches such as Helmholtz resonators, wall treatments using porous materials, and traditional acoustic liners. These techniques have the advantage of being fully passive, in other words they do not inject energy into the system and therefore do not generate additional noise. However, their main drawback is their limited effectiveness at low frequencies, where high absorption typically requires large volumes or thick treatments.

There are also active approaches, with the main representative being active noise control (ANC). In these techniques, microphones are placed in a cavity to measure the sound field and one or more loudspeakers generate anti-noise in order to cancel the unwanted sound in targeted regions. While these solutions can be effective at low frequencies, they typically



require an accurate knowledge or identification of the acoustic cavity where they are applied. Additionally, since they inject energy into the system, they are more prone to instabilities. In this context, Electroacoustic absorbers (EAs) present themselves as an interesting bridge between passive and active techniques.

EAs are systems composed of a loudspeaker membrane and four collocated microphones (Fig. 1). They are capable of modifying the acoustic impedance at their surface, Z_{at} . This is achieved by using the pressure measurements $P(s)$ from the microphones to impose a target velocity $V(s)$. The earliest pairing of a loudspeaker with a collocated microphone can

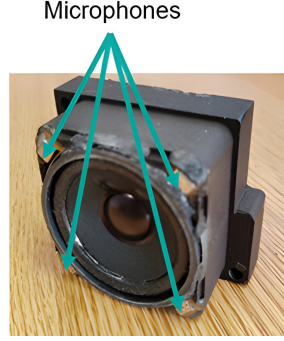


Figure 1. Electroacoustic absorber

be traced back to 1953, when Olson and May (1953) introduced the so-called “electronic sound absorber.” In this work, the main goal of the system was to minimize the sound pressure near the surface of the loudspeaker. Although this is not explicitly stated in the paper, it can be interpreted that the system effectively targets a near zero acoustic impedance at the microphones vicinity.

Using Olson’s work as inspiration, Guicking and Lorenz (1984) investigated low-frequency absorption by coupling a loudspeaker to a porous layer and actively enforcing a backing nearly zero pressure in the air gap behind the porous layer, thereby realizing an active equivalent of a $\lambda/4$ resonance absorber, thereby avoiding the need for a physically deep cavity.

An explicit reference to active impedance control was made later by Orduña-Bustamante and Nelson (1992). In this case, the system was implemented using a two-microphone configuration inside a duct with the loudspeaker cell at its termination. Subsequently, Furstoss *et al.* (1997) employed a system composed of near-collocated accelerometers and microphones to achieve impedance control.

Finally, Collet *et al.* (2009) evaluated the use of impedance control along a lined surface equipped with multiple loudspeakers. The focus of this work was to demonstrate the capability of strongly modifying the behavior of a waveguide. Subsequently, Lissek (2009) presented an electrically shunted loudspeaker in which an electrical load is inserted into the loudspeaker to modify its electrical dynamics, thereby altering its acoustic impedance. Later, a microphone-based feedforward control strategy employing a model inversion controller was proposed to target a user-defined impedance (Rivet *et al.*, 2017). These approaches have also been employed for noise attenuation inside acoustic cavities (De Bono *et al.*, 2024) and in waveguides (De Bono *et al.*, 2025a,b).

The stability of the control algorithm is affected by uncertainties such as temperature variations (Ferreira *et al.*, 2026) and time delay (De Bono *et al.*, 2022). Considering the limitations imposed by time delay, this work investigates the potential of using reinforcement learning techniques to achieve the absorption of single harmonics while maintaining system stability, as a limiting factor of this study only an analytical evaluation of the performance will be evaluated.

This paper is divided into five main sections: Section 2 presents the stability analysis of the model inversion control and introduces the disk margin criterion; Section 3 describes the reinforcement learning framework, its fundamental concepts and makes a brief explanation on the theory of the technique used on this work; Given the informations presented in Section 3, Section 4 details the design process behind the Reward function utilized on this case and present the main hyperparameters used during the training, including the definition of the state, action space and reward function; Section 5 discusses the obtained results from the training; finally, Section 6 summarizes the main conclusions and possible next steps of this work.

2. STABILITY IN A MODEL INVERSION CONTROL

The model inversion control algorithm is a control strategy designed to impose a desired acoustic impedance on the electroacoustic absorber by acting on its electrical input. By inverting the dynamics of the uncontrolled system, it becomes possible to drive the absorber toward a prescribed target impedance. Accordingly, the controller is defined by the transfer function $H(s)$ given by:

$$H(s) = \frac{I(s)}{P(s)} = \frac{S_d}{Bl} \left(1 - \frac{Z_{a0}(s)}{Z_{at}(s)} \right) \quad (1)$$

where $I(s)$ is the electrical current sent to the EA (in A), S_d is the effective diaphragm area (in m²), Bl is the electromechanical force factor (in N/A), s is $j\omega$ and $Z_{a0}(s)$ is the uncontrolled acoustic impedance of the resonator (in Pa·s/m).

The target impedance $Z_{at}(s)$ is defined as a single-degree-of-freedom impedance given by follows the single-piston approximation:

$$Z_{at}(s) = \mu_m M_{a0}s + \mu_r \rho_0 c_0 + \frac{\mu_k K_{a0}}{s}, \quad (2)$$

where M_{a0} is the acoustic mass of the resonator (in kg/m²), K_{a0} is the acoustic stiffness (in N/m³), ρ_0 is the air density (in kg/m³), c_0 is the speed of sound (in m/s), and μ_m , μ_k , and μ_r are the mass, stiffness, and resistance control variables, respectively. By adjusting μ_m , μ_k , and μ_r , different impedance configurations can be achieved.

As with any control algorithm, the system is subject to time delay τ (s). In this case, the achieved impedance under delay, denoted $Z_d(s)$, can be approximated by:

$$Z_d(s) = \frac{Z_{a0}(s)}{1 - \frac{Bl}{S_d} H(s) e^{-s\tau}}. \quad (3)$$

From Z_{at} , and consequently from Z_d as well, we can extract respectively the targeted reflection coefficient R and absorption coefficient α from Eq. (4) and Eq. (5). R is a measure that shows how much of the waves are reflected per frequency and more important in the case of this application α shows the sound absorption profile of the EA:

$$R = \frac{Z_{at} - \rho_0 c_0}{Z_{at} + \rho_0 c_0}, \quad (4)$$

$$\alpha = 1 - |R|^2. \quad (5)$$

On Fig. 2 it is possible to observe the effect of the time-delay in the absorption coefficient while changing μ_m (in this case $\mu_k = 1$ and $\mu_r = 1$). Observe that as the value of μ_m decreases, the frequency in which $\alpha = 1$ increases. Also observe that, at the same time, a loss of passivity ($\alpha < 0$) that was taking place around 600 Hz starts to happen after 2000 Hz. This means, theoretically, that a localized emission of noise that was taking place around 600 Hz is becomes a broadband emission at frequencies superior to 2000 Hz.

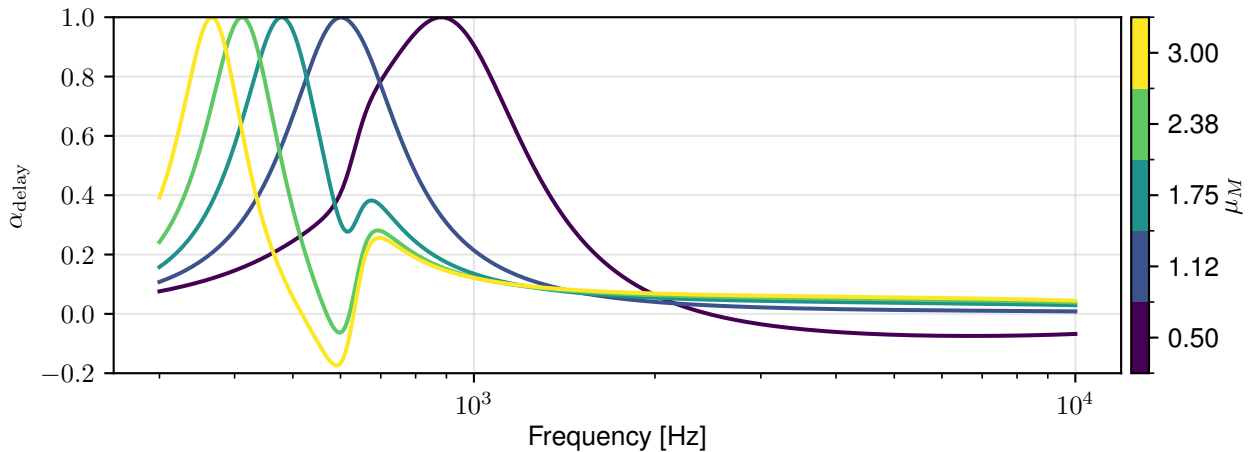


Figure 2. Absorption coefficient when varying μ_m

But this loss does not imply an unstable system, since at higher frequencies the model used to create the model-inversion control loses its validity. Consequently, the impedance set by the model does not correspond to the actually experimentally achieved impedance (even when considering the time-delay effect), making $H(s)$ insufficient to guarantee system stability. Therefore, a closer evaluation of the interaction between $H(s)$ and the plant in which it is inserted is required.

This investigation was carried out by Volery and Lissek (2023) in the case of a mixed feedforward–feedback controller; however, the same methodology can be transposed to the model inversion control case.

In the present work, the control system is therefore formulated as a feedback configuration, as shown in Fig. 3:

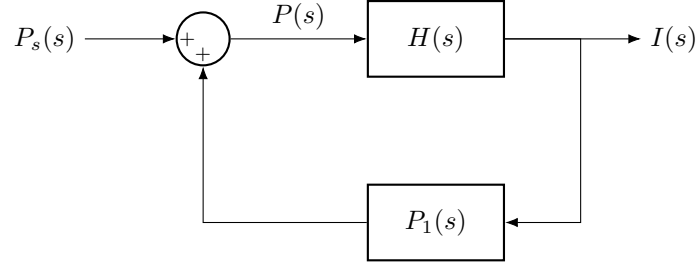


Figure 3. Block diagram of model inversion transfer function

where P_s denotes the input pressure (in Pa) coming from the source and $P_1(s)$ indicates the plant transfer function defined as:

$$P_1(s) = \frac{Bl}{S_d} \frac{e^{-s\tau}}{1 - \frac{Z_{a0}(s)}{Z_{rad}(s)}}. \quad (6)$$

In with $Z_{rad}(s)$ is the radiated impedance, which for a cavity of length L in the one-dimensional case is given by:

$$Z_{rad}(s) = -\rho_0 c_0 \frac{e^{sL/c_0} + R_{term} e^{-sL/c_0}}{e^{sL/c_0} - R_{term} e^{-sL/c_0}}. \quad (7)$$

For this work, P_1 is experimentally measured following the procedure established by Volery (2023).

Parting from P_1 it is possible to obtain the open-loop transfer function of the system, $G(s)$:

$$G(s) = H(s)P_1(s). \quad (8)$$

From $G(s)$ stability can be assessed through classical analysis of the disk margin of the system using the Nyquist diagram (Barrett, 1980). The disk margin is defined as the minimum distance between the Nyquist curve and the critical point $(-1, 0)$ and is given by:

$$D_m = \min_s |G(s) + 1|. \quad (9)$$

Figure 4 provides a graphical representation of the disk margin associated with $G(s)$ for the electroacoustic absorber. The Nyquist curve of $G(s)$ is shown, and the disk margin D_m is highlighted in red, corresponding to the minimum distance between the curve and the critical point $(-1, 0)$ for this specific $G(s)$, as it gets closer to $(-1, 0)$ the system tends to become unstable.

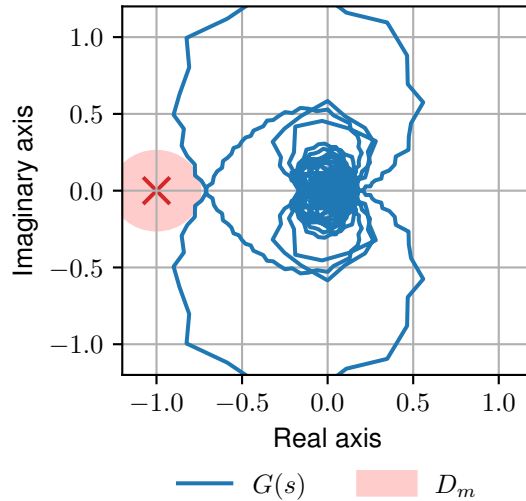


Figure 4. $G(s)$ Nyquist curve with obtained D_m

3. THE DEEP DETERMINISTIC POLICY GRADIENT ALGORITHM

Reinforcement learning is a family of machine learning algorithms that optimizes a policy over time through repeated interactions between an agent and an environment, guided by a reward function. At each step, the agent observes the current state S and selects an action a according to its policy; this action alters the environment, yielding a reward R and a new state S' . The process then repeats from S' , with the agent choosing the next action a' , with the objective of maximizing the cumulative reward over time. Figure 5 illustrates this interaction loop.

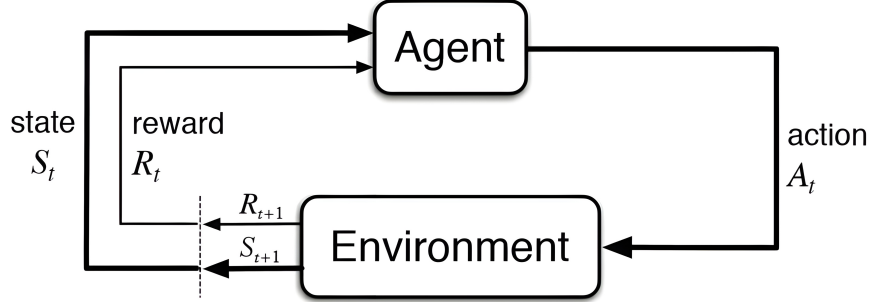


Figure 5. Reinforcement learning schematic (Sutton and Barto, 2018)

Most modern reinforcement learning techniques focus on the optimization of the so-called value function $Q(S, a)$ (see Eq. (10)), which satisfies the Bellman equation (Bellman, 1957),

$$Q(S, a) = R + \gamma Q(S', a'). \quad (10)$$

The value function represents the expected cumulative reward obtained by taking action a in state S and following the learned policy thereafter. The discount factor γ determines the importance of future rewards. During training, many algorithms construct an approximation of this function and, by maximizing it, converge toward an optimal policy.

In this work, the Deep Deterministic Policy Gradient (DDPG) algorithm (Lillicrap *et al.*, 2016) is employed. DDPG is an actor-critic algorithm that uses neural networks to approximate both the policy and the value function. In this framework, the critic learns an approximation of the environment's value function, while the actor uses this learned model to optimize its policy by selecting actions that maximize the predicted value function. This optimization is performed using the Deterministic Policy Gradient (DPG) (Silver *et al.*, 2014) of the critic, in which an ascending gradient step is applied with respect to the actor parameters θ . The loss functions of the critic ($\mathcal{L}_c$) and the actor ($\mathcal{L}_a$) are given by:

$$\begin{cases} \mathcal{L}_c = \mathbb{E} \left[(Q(S, a) - (R + \gamma Q(S', a')))^2 \right], \\ \mathcal{L}_a = -\nabla_{\theta} [Q(S, a)] = -\nabla_a Q(S, a) \nabla_{\theta} a(S). \end{cases} \quad (11)$$

In DDPG, four neural networks are employed: two actor networks and two critic networks. This four-network architecture is derived from the two-network Deep Q-Learning architecture introduced by Mnih *et al.* (2015). It was originally proposed as a mechanism to improve training convergence in Deep Reinforcement Learning (reinforcement learning techniques based on deep neural networks).

In the original Deep Q-Learning formulation, a target network is used as a frozen reference for the main network during the optimization process. This provides the main network with a stable objective to approximate. Every n iterations, the target network parameters are updated using the parameters of the main network. By introducing this delayed update mechanism, the algorithm improves the convergence behavior of the estimated $Q(S, a)$.

In the case of DDPG, instead of updating the target network every n iterations, the target networks are updated at every iteration using a soft update mechanism. This means that the target network parameters are slowly adjusted toward the parameters of the main network. The update process is regulated by a soft update coefficient.

Additionally, as a general practice in reinforcement learning algorithms, an exploration policy is adopted throughout training. This exploration mechanism is typically implemented by introducing noise into the action selection process. One common example is the ϵ -greedy policy, in which a probability is assigned for the agent to take a random action during training. This exploration allows the algorithm to better generalize and avoid convergence to local solutions. In this study, the exploration policy is implemented by adding a stochastic noise to the actions during the training process.

4. TRAINING SETUP

In this study, the problem of single-harmonic absorption is used as a case example to assess the potential of reinforcement learning. At each training step, the agent receives as observation a desired frequency f_d in the range 300-3000 Hz

(observation space). Based on f_d , it selects continuous values of μ_m , μ_k , and μ_r within $[0.1, 5.0]$ (action space). These parameters modify the achieved impedance Z_d , which in turn changes the obtained absorption coefficient α . The agent's objective is therefore to select μ_m , μ_k , and μ_r so as to maximize α at the specified frequency f_d .

To enable the agent to correctly achieve this objective, a two-component reward function is designed:

$$R = P_i S_i, \quad (12)$$

where P_i represents the performance index of the control algorithm and S_i represents the stability index. The performance index P_i is defined as the absorption coefficient evaluated at f_d :

$$P_i = \alpha(\mu_m, \mu_k, \mu_r, f_d). \quad (13)$$

Finally, the stability index S_i ensures that only stable solutions are rewarded. Stability is assessed using the previously introduced disk margin criterion. A minimum acceptable disk margin, denoted by D_{lim} , is defined. The stability index is then determined by Eq. (14). In this study, D_{lim} is set to 0.2:

$$S_i = \begin{cases} 1, & \text{if } D_m \geq D_{lim}, \\ 0, & \text{if } D_m < D_{lim}. \end{cases} \quad (14)$$

Having defined the reward, the algorithm was trained over 100,000 episodes. Each episode consists of a single step, which characterizes this problem as a bandit problem (i.e., a single-step reinforcement learning problem). Standard network architectures and hyperparameters were used for training. They are reported in Tab. 1, Tab. 2, and Tab. 3.

Table 1. Architecture of the DDPG actor network.

Layer	Layer Type	Input Dim	Output Dim	Activation Function
1	Input	1	1	–
2	Fully Connected Layer 1	1	400	ReLU
3	Fully Connected Layer 2	400	300	ReLU
4	Output Layer	300	3	Tanh

Table 2. Architecture of the DDPG critic network.

Layer	Layer Type	Input Dim	Output Dim	Activation Function
1	Input (State + Action Concatenation)	4	4	–
2	Fully Connected Layer 1	4	400	ReLU
3	Fully Connected Layer 2	400	300	ReLU
4	Output Layer	300	1	–

Table 3. DDPG training hyperparameters.

Hyperparameter	Value
Learning rate	1×10^{-3}
Batch size	64
Soft update coefficient	$\tau = 0.005$
Discount factor	$\gamma = 0$
Exploration noise type	Gaussian
Noise mean	$\mu = 0$
Noise standard deviation	$\sigma = 0.01$

The optimization of the training hyperparameters is not the focus of this study. However, for problem-definition purposes, one important modification was introduced: the discount factor γ was set to 0. This choice was made to explicitly formulate the problem as a bandit problem, meaning that the algorithm targets an instantaneous reward rather than the optimization of cumulative reward over time. Under this condition, the value function reduces to the immediate reward:

$$Q(S, a) = R. \quad (15)$$

5. Training Results

The analysis of the results is divided into two subsections. The first is concerned with the evaluation of the algorithm strictly from a reinforcement learning perspective. In this part, the reward function is analyzed, the general convergence behavior of the algorithm is assessed, and the main training dynamics are discussed.

The second subsection focuses on the acoustic performance of the algorithm. The operational range is analyzed, and comments are made regarding the stability of the solution when the algorithm operates both inside and outside this range.

5.1 Reward Analysis

The reward curve obtained during training is shown in Fig. 6. Two curves are presented: in blue, the episodic reward during training, and in red, the mean average reward. The episodic reward, as the name implies, shows the reward obtained at each training iteration, while the mean average reward represents the moving average of the episodic reward.

It can be observed that the reward converges after approximately 15000 training steps. However, a significant variability remains throughout the training process. This variability can be explained by two main factors. First, when the algorithm targets frequencies outside its feasible operational range, the reward is reduced. Second, and most important, due to the presence of exploration noise, the algorithm does not necessarily execute exactly the action predicted by the actor network, in particular this second reason explains the presence of steps in with the reward was 0, since the algorithm operates in the limit of the stability the presence of noise leads it to take actions that may lead to an unstable system.

Although this behavior may initially appear detrimental to the training process, it actually contributes to improved generalization. For this reason, the analysis of the mean average reward is essential, as it filters the variability introduced by the exploration noise and better reflects the convergence trend of the algorithm.

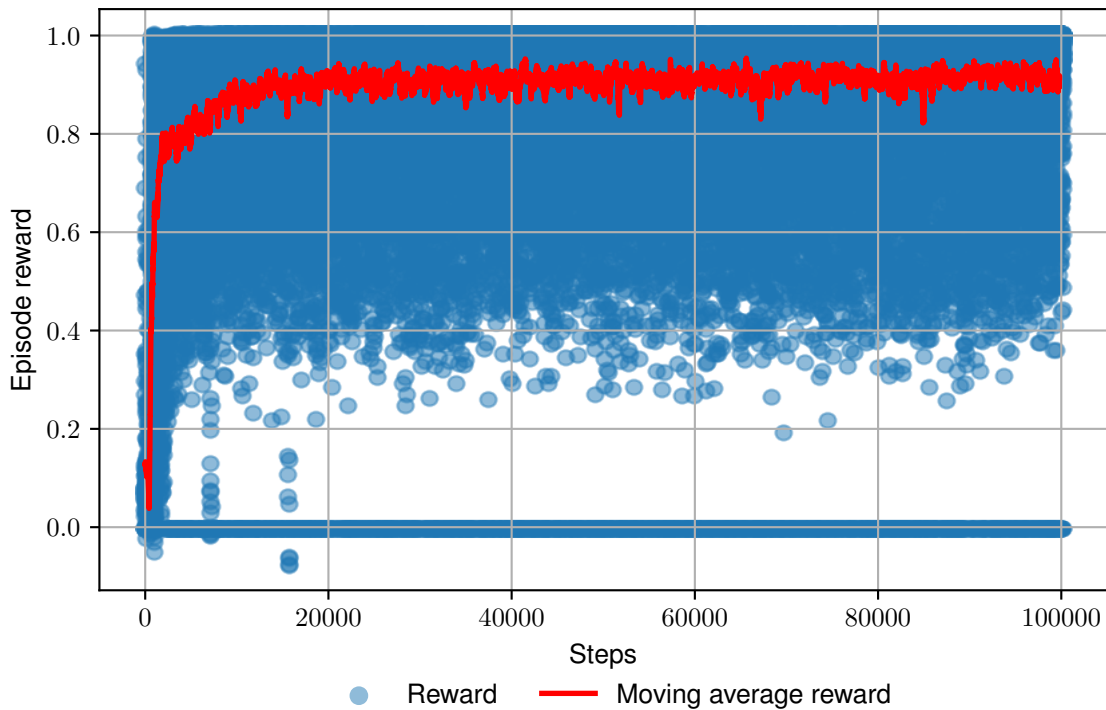


Figure 6. Reward curve obtained during training. Blue: episodic reward. Red: moving average reward.

5.2 Acoustic Analysis

For the acoustic performance analysis, the trained model was evaluated without exploration noise over the frequency range from 300 Hz to 3000 Hz. During this evaluation the value of α , obtained analytically using Eq.5, for each desired frequency f_d was recorded together with the corresponding stability margins.

Based on these results, the operational range of the control algorithm was then determined. The operational range is defined as the set of target frequencies f_d for which the algorithm is capable of achieving $\alpha \geq 0.9$, corresponding to at least 90% absorption of the acoustic energy. As indicated by the light blue region in Fig. 7, the operational range extends from 300 Hz to 2600 Hz. Consistently with this definition, the dark blue curve, which represents the value of α obtained for each target frequency, remains above 0.9 throughout this interval.

To further illustrate these results, three absorption coefficient curves are also shown in Fig. 7. Each of these curves represents the full absorption coefficient obtained when the algorithm is tasked with maximizing absorption at a specific target frequency f_d . The orange and green curves correspond to $f_d = 300$ Hz and $f_d = 1500$ Hz, respectively, both of which lie within the operational range. In both cases, the algorithm successfully achieves $\alpha \geq 0.9$.

By contrast, the red curve corresponds to the case $f_d = 3000$ Hz, which lies outside the operational range. In this case, the maximum achieved absorption is approximately $\alpha \approx 0.55$. This limitation arises because the algorithm must also satisfy the stability constraint imposed by the reward function.

This stability seeking behavior is further illustrated in Fig. 8, where the Nyquist diagrams of the open-loop transfer function are shown for the three cases described above. It can be observed that, regardless of the target frequency, the open-loop transfer function does not enter the light red region defined by D_{lim} . This comportment is consistent with the design of the reward function: the algorithm avoids selecting values of μ_m , μ_k , and μ_r that would violate the stability constraint, and instead selects the parameters that yield the highest achievable α while preserving stability.

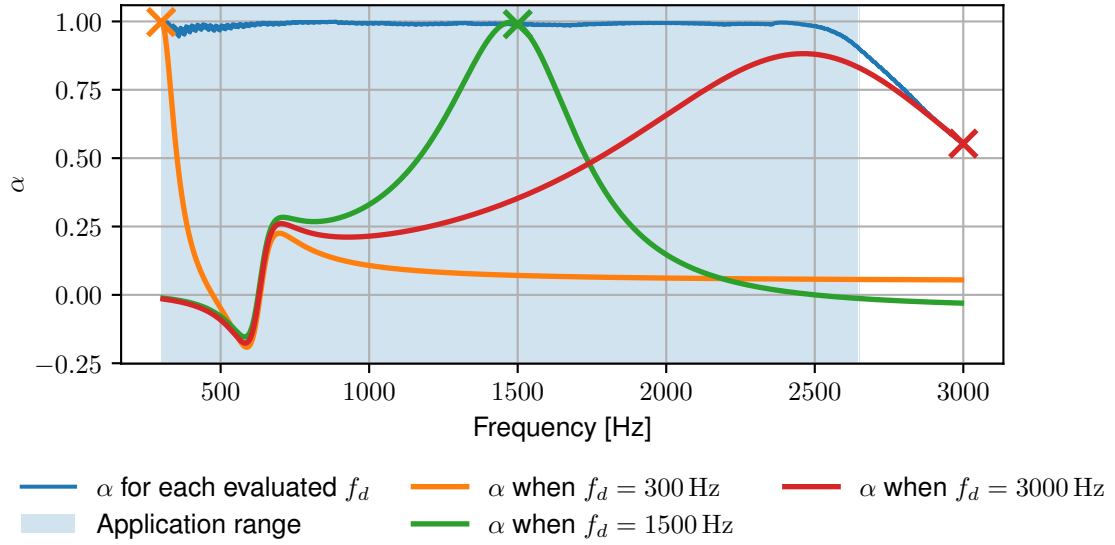


Figure 7. The algorithm's application range , together with the absorption coefficient obtained at each value of f_d . For visualization purposes, three complete absorption coefficient curves corresponding to $f_d = 300$ Hz, 1500 Hz, and 3000 Hz are also plotted.

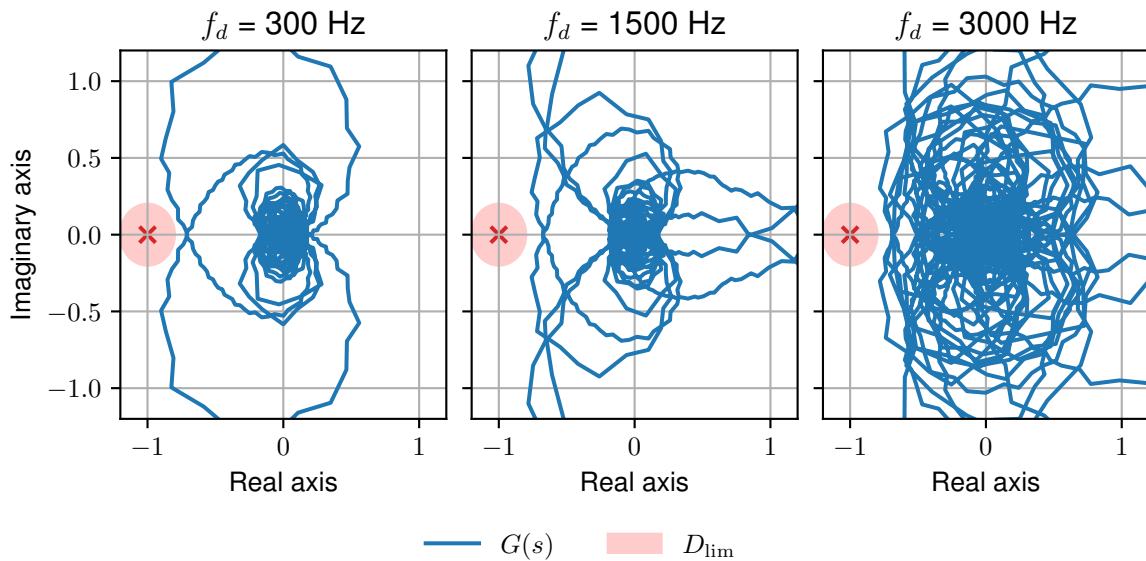


Figure 8. Nyquist diagrams evaluated at $f_d = 300$ Hz, 1500 Hz, and 3000 Hz. In all cases, the stability constraint defined by D_{lim} is respected.

6. CONCLUSIONS

The algorithm was capable of maintaining the defined stability criterion throughout the training range. It was able to correctly target absorption at certain frequencies and, when it was not able to achieve the desired absorption, it adjusted its parameters in order to obtain the maximum possible absorption while preserving stability.

Given these results a natural next step would be to proceed to an experimental training/validation of this work. Nevertheless, the algorithm required more than 15000 iterations to converge, a number that would make training in an experimental environment unviable. Additionally, even though during evaluation the algorithm was able to obey the restrictions imposed during the training process, during training it was subjected to an exploration policy and therefore it constantly violated the stability region assigned to it. This behavior could prove troublesome for experimental training. In this context, training in a model would be more desirable.

One possible avenue for improvement would be to explore the application of model-based reinforcement learning techniques. In this approach, a surrogate model of the environment would be used as a part of the algorithm and through a process of dynamic exploration the agent would be capable of achieving an optimal behavior.

A longer-term prospect for this research would be to explore methods for the system to adapt to unforeseen changes in the environment. In practical applications, uncertainties such as temperature variations and aging of the loudspeaker can modify the dynamics of the EA over time. In this case, it would be desirable for the algorithm to identify these changes and based on this adapt its behavior to different scenarios.

Lastly, the application of multiple cells in a multi-agent architecture should also be explored. By combining several EAs in a distributed configuration, it would be possible to control multiple entities independently and, by cooperation or competition, they could achieve an emerging behavior on the system. This approach could also account for variations in the structure, for example structures composed by different types of cells. Finally, it could allow the system to adapt to changes in the environment, for example a broken loudspeaker and non-global temperature changes.

7. REFERENCES

- Barrett, M.F., 1980. "Conservatism with robustness tests for linear feedback control systems". In *Proceedings of the 19th IEEE Conference on Decision and Control*. pp. 885–890.
- Bellman, R., 1957. *Dynamic Programming*. Dover Publications. ISBN 9780486428093.
- Collet, M., David, P. and Berthillier, M., 2009. "Active acoustical impedance using distributed electrodynamical transducers". *The Journal of the Acoustical Society of America*, Vol. 125, No. 2, pp. 882–894. doi:10.1121/1.3026329. URL <https://doi.org/10.1121/1.3026329>.
- De Bono, E., Collet, M., Matten, G., Karkar, S., Lissek, H., Ouisse, M., Billon, K., Laurence, T. and Volery, M., 2022. "Effect of time delay on the impedance control of a pressure-based, current-driven electroacoustic absorber". *Journal of Sound and Vibration*, Vol. 537, p. 117201. doi:10.1016/j.jsv.2022.117201.
- De Bono, E., Collet, M., Ouisse, M., Salze, E., Volery, M., Lissek, H. and Mardjono, J., 2025a. "The advection boundary law in presence of mean flow and plane wave excitation: Passivity, non-reciprocity and enhanced noise transmission attenuation". *Journal of Sound and Vibration*, Vol. 618, p. 119293. doi:10.1016/j.jsv.2025.119293.
- De Bono, E., Ponticelli, D., Rosa, S.D., Petrone, G., Ouisse, M., Sadoulet-Reboul, E. and Teloli, R., 2024. "Active impedance control optimization for attenuation of acoustic cavity modes". In *Proceedings of the 30th AIAA/CEAS Aeroacoustics Conference*. Rome, Italy. doi:10.2514/6.2024-3195. AIAA Paper 2024-3195.
- De Bono, E., Salze, E., Collet, M., Gillet, M., Ouisse, M., Volery, M., Lissek, H. and Mardjono, J., 2025b. "Experimental assessment of a programmable electroacoustic liner in a representative turbofan facility". *Applied Acoustics*, Vol. 240, p. 110896. doi:10.1016/j.apacoust.2025.110896.
- Ferreira, L., de O Teloli, R., De Bono, E. and Ouisse, M., 2026. "Effect of the temperature on the impedance control of pressure-based, current-driven electroacoustic absorbers: Addressing the loss of passivity using a viscoelastic material model". *Journal of Sound and Vibration*, Vol. 621, p. 119468. doi:10.1016/j.jsv.2025.119468.
- Furstoss, M., Thenail, D. and Galland, M.A., 1997. "Surface impedance control for sound absorption: direct and hybrid passive/active strategies". *Journal of Sound and Vibration*, Vol. 203, No. 2, pp. 219–236. doi:10.1006/jsvi.1996.0905. URL <https://doi.org/10.1006/jsvi.1996.0905>.
- Guicking, D. and Lorenz, E., 1984. "An active sound absorber with porous plate". *Journal of Vibration, Acoustics, Stress, and Reliability in Design*, Vol. 106, No. 3, pp. 389–392. doi:10.1115/1.3269206. URL <https://doi.org/10.1115/1.3269206>.
- Lillicrap, T.P., Hunt, J.J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D. and Wierstra, D., 2016. "Continuous control with deep reinforcement learning". In *International Conference on Learning Representations (ICLR)*.
- Lissek, H., 2009. "Electroacoustic shunt loudspeaker techniques for variable impedance engine nacelle liners". In *Proceedings of Inter-Noise 2009*. URL <https://infoscience.epfl.ch/handle/20.500.14299/40131>.
- Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg,

- S. and Hassabis, D., 2015. “Human-level control through deep reinforcement learning”. *Nature*, Vol. 518, No. 7540, pp. 529–533. doi:10.1038/nature14236.
- Olson, H.F. and May, E.G., 1953. “Electronic sound absorber”. *The Journal of the Acoustical Society of America*, Vol. 25, No. 6, pp. 1130–1136. doi:10.1121/1.1907249. URL <https://doi.org/10.1121/1.1907249>.
- Orduña-Bustamante, F. and Nelson, P.A., 1992. “An adaptive controller for the active absorption of sound”. *The Journal of the Acoustical Society of America*, Vol. 91, No. 5, pp. 2740–2747. doi:10.1121/1.403779. URL <https://doi.org/10.1121/1.403779>.
- Rivet, E., Karkar, S. and Lissek, H., 2017. “Broadband low-frequency electroacoustic absorbers through hybrid sensor-/shunt-based impedance control”. *IEEE Transactions on Control Systems Technology*, Vol. 25, No. 1, pp. 1–10.
- Silver, D., Lever, G., Heess, N., Degris, T., Wierstra, D. and Riedmiller, M., 2014. “Deterministic policy gradient algorithms”. In *Proceedings of the 31st International Conference on Machine Learning (ICML)*. pp. 387–395.
- Sutton, R.S. and Barto, A.G., 2018. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, 2nd edition.
- Volery, M. and Lissek, H., 2023. “Stability limits of an electroacoustic absorber”. In *Proceedings of the 10th Convention of the European Acoustics Association - Forum Acusticum 2023*. Turin, Italy, pp. 3651–3654. doi:10.61782/fa.2023.0703.
- Volery, M., 2023. *Robust tunable acoustic impedance control on electroacoustic resonators for aircraft noise reduction*. Ph.D. thesis, EPFL, Lausanne. doi:10.5075/epfl-thesis-9833. URL <https://infoscience.epfl.ch/handle/20.500.14299/200049>.

8. RESPONSIBILITY NOTICE

The authors are solely responsible for the printed material included in this paper.