Semi-Active Control of a Shear Building based on Reinforcement Learning: Robustness to measurement noise and model error

Aleksandra Jedlińska*, Dominik Pisarski, Grzegorz Mikułowski,
Bartłomiej Błachowski, Łukasz Jankowski
0009-0003-0644-0760, 0000-0002-0515-3298, 0000-0003-4215-8650, 0000-
0001-6021-0374, 0000-0002-9773-0688
Institute of Fundamental Technological Research (IPPT PAN),
Polish Academy of Sciences 02-106 Warsaw, Poland
Email: ajedlins@ippt.pan.pl | ljank@ippt.pan.pl

Abstract—**This paper considers structural control by reinforcement learning. The aim is to mitigate vibrations of a shear building subjected to an earthquake-like excitation and fitted with a semi-active tuned mass damper (TMD). The control force is coupled with the structural response, making the problem intrinsically nonlinear and challenging to solve using classical methods. Structural control by reinforcement learning has not been extensively explored yet. Here, Deep-Q-Learning is used, which apprivimates the Q-function with a neural network and optimizes initially random control sequences through interaction with the controlled system. For safety reasons, training must be performed using an inevitably inexact numerical model instead of the real structure. It is thus crucial to assess the robustness of the control with respect to measurement noise and model errors. It is verified to significantly outperform an optimally tuned conventional TMD, and the key outcome is the high robustness to measurement noise and model error.**

Index Terms—**structural control, semi-active control, reinforcement learning, tuned mass damper (TMD).**

I. INTRODUCTION

In this paper, a novel control strategy for reducing structural vibrations in shear-type building structures under seismic excitation is presented and assessed. To achieve this, machine learning techniques, specifically reinforcement learning (RL), were customized, developed, and applied. Structural vibrations in engineering structures can have a negative impact on structural condition and operation, and they can negatively impact structural integrity. Various approaches have been developed to mitigate these effects, including passive, active, and semi-active control methods [1], [2]. The semi-active methods are appealing, since they do not require significant power sources and can be designed to be failure-safe. However, the control forces are coupled with the structural response, which leads to formulations that are challenging to be solved using classical methods [3], [4]. This paper focuses on semi-active control through the use of a semi-active tuned mass damper (TMD). The TMD is a classical device used to mitigate structural vibrations by adding a secondary mass that opposes the motion of the main structure [5], [6]. The semi-active TMD applied here is controllable through a switchable level of viscous damping.

The main aim of this contribution is to test the application potential of reinforcement learning (RL) in semi-active structural control, and in particular, the robustness of the trained agent to measurement noise and structural errors. This is a crucial problem for potential practical applications in civil engineering, since for safety reasons the RL agent must be trained using a numerical model instead of the physical target structure. The structure investigated here is an 11-story shear-type building equipped with a semi-active TMD. The structure is modeled using the finite element (FE) method, and the specific parameters of the models are taken from literature [7]. The TMD is controlled by switching its viscous damping coefficient in an on/off manner (bang–bang), as suggested by the Pontryagin minimum principle [8]. The structure is subjected to an earthquake-like random base excitation. A Deep Q Learning (DQN) algorithm is applied. The trained RL agent reduces the structural vibrations effectively and to a greater extent than a conventional tuned mass damper. Importantly, the contribution demonstrates and evaluates *also* the robustness of the trained agent with respect to measurement noise and model error.

II. REINFORCEMENT LEARNING – THE TECHNIQUE AND SYSTEM ARCHITECTURE

Reinforcement learning is a set of machine learning techniques that aim to teach an agent to determine the most effective actions by engaging in trial-and-error interactions with its environment. During the process the agent receives feedback in the form of rewards or punishments, which it uses to enhance its decision-making abilities over time. This research investigates the capability of reinforcement learning (RL) to enhance semi-active structural control. Unlike supervised learning that depends on optimal control sequences, which are often unknown in semi-active control, and unlike unsupervised learning, which solely relies on exploring input data, RL enables learning from interactions and seems to be well-tailored to the needs of structural control. However, despite the large successes of RL in mastering other complex tasks [9], including control-like problems [10]-[13], it is still a novel and very scarcely explored approach in structural control with only a handful of publications [14]-[16].

Technical track: Software Engineering for
Cyber-Physical Systems

In this study, the reinforcement learning (RL) agent employs a dense artificial neural network (ANN) to learn and encode the Q-function. The ANN is implemented in the Python programming language using two popular open-source libraries, TensorFlow and Keras. TensorFlow is a low-level library used for building and training machine learning models, while Keras is a high-level API that simplifies the process of building neural networks. The ANN used in this study consists of six hidden sequential dense layers, each with 40 neurons. The input layer provides the network with measurements of structural response, while the output layer consists of two neurons corresponding to the possible states of the control signal. The activation function used in the neural network is rectified linear unit (ReLU) [17].

III. SHEAR BUILDING AND EXCITATION

The structure analyzed in this study is a shear-type building consisting of 11 stories with a semi-active tuned mass damper (TMD) attached to the top story (Fig. 1). The TMD is a well-known classical engineering device that comprises a mass, spring, and viscous damper and is widely used to reduce vibrations in structures subjected to external excitations, such as earthquakes [5]. Such a setup results in a total of twelve degrees of freedom (DOFs) which correspond to each of the eleven stories and the TMD. The equation of motion for the building model under seismic excitation can be expressed as:

$$[M]\{\ddot{u}\} + [C]\{\dot{u}\} + [K]\{u\} = -[M]\{r\}a(t) \quad (1)$$

The vector $\{u\}$ has 12 rows and represents the absolute displacements of each DOF, while the vector $\{r\}$ also has 12 rows and represents the displacement resulting from unit horizontal ground displacement for each DOF. The ground acceleration is denoted by $a(t)$, while the matrices $[M]$, $[C]$, and $[K]$ are 12×12 in dimension, and represent the mass, damping, and stiffness of the structure, respectively. The material damping model is assumed, and the damping matrix $[C]$ is proportional to the stiffness matrix with the proportionality coefficient chosen to achieve 2% critical damping for the first mode of vibration of the structure without the TMD. The control directly affects the entries in $[C]$ that correspond to the damping of the TMD, see Fig. 1, by switching it between zero and a large value. The mass matrix $[M]$ is diagonal. The masses of each story and the TMD are thus listed on the diagonal of the mass matrix, assuming lumped masses at each floor level. Building specifications, including the number of stories, their masses, and stiffnesses, are based on the literature data [7]. The first undamped natural frequency is 0.89 Hz for the building with the TMD and 1.05 Hz for the building without the TMD.

In this study, an effort has been made to safeguard the RL agent from acquiring a limited response pattern conditioned on a particular collection of ground movements. This restriction is essential to avoid overfitting, a common issue in supervised learning. To address this concern, the seismic load $a(t)$ is assumed to be the white Gaussian noise. Consequently,

it is generated afresh for every training and evaluation episode, guaranteeing that the proposed control system is exposed to diverse ground motions without any bias towards specific patterns [7].
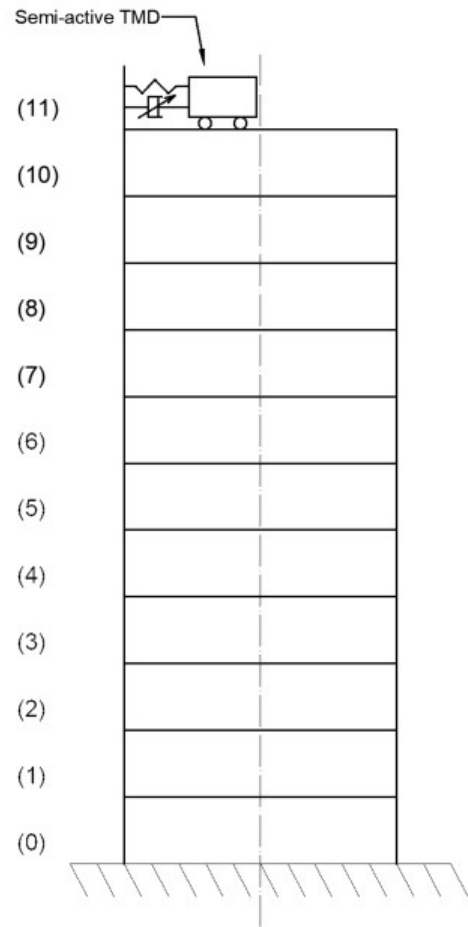

Fig. 1. The investigated 11-DOF structure with a semi-actively controlled TMD placed on the top level

IV. RL TRAINING

The state of the RL environment employed for training and control purposes is based on linearly transformed full structural state vector, and it is comprised of the relative displacements and velocities between the ground, subsequent floors, and the TMD. Such a choice is practical, as the relative inter-story displacements and velocities are relatively easy to be measured in a real setting.

The training proceeds in episodes. Each episode consists of 1000 RL steps of 25 ms each, and it corresponds to about 25 periods of the fundamental structural vibration. For fidelity of structural response simulation, each RL step is internally further subdivided into 5 simulation steps, each of 5 ms.

The aim of the control is to reduce the oscillations experienced by the highest floor of the structure. For structural control purposes, the control efficiency is usually assessed using the root mean square (RMS) of the displacements in each ep-

isode. Consequently, the agent's training is based on the rewards it receives in each step of the interaction episodes, and the rewards are evaluated using the displacement level of the top floor. The maximum reward of 1 is assigned when the displacement is zero. The rewards are used to update the agent's Q-function, allowing it to improve its performance in future episodes.

The total reward signal at the end of each episode reflects the agent's performance and the distance from the equilibrium point over all time steps. Fig. 2 shows the total reward per training episode, together with its EMA50, which increases as the agent learns. The value of 1000 denotes a perfectly stationary top floor, and the chart includes the effect of a 10% exploration rate (10% of actions, on average, is selected randomly to ensure ongoing exploration of the action space).
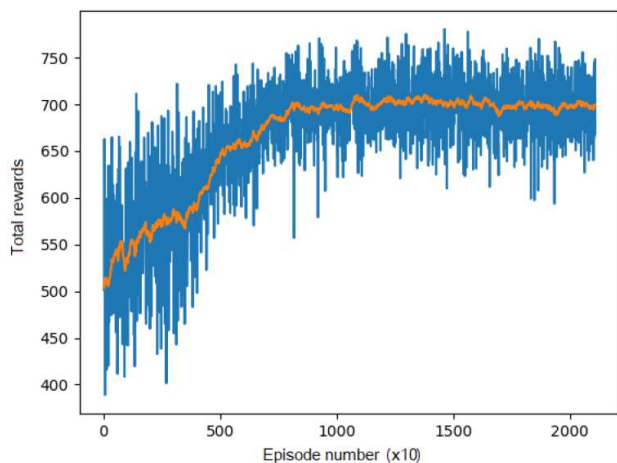

Fig 2. Total rewards per training episode (blue line) and its EMA50 (orange line)

V. ROBUSTNESS TO MEASUREMENT NOISE AND MODEL ERROR

The intended ultimate application scenario involves a real physical environment (building structure) rather than just its idealized mathematical model. There are two main factors that inevitably differentiate a physical structure from its numerical model: 1) measurement noises overlaid on signals from physical sensors, and 2) model errors that represent the modeling inaccuracies. These factors can negatively affect the control efficiency applied by an RL agent trained using an idealized environment.

The first test involves applying simulated measurement noise to the agent's observations, which is modeled as a Gaussian white noise and added to the input to the neural network (sensor measurements). The test examines increasing larger levels of noise, which is quantified in the signal RMS terms (noise standard deviation related to the RMS of the original sensor signal). The control effectiveness is assessed in terms of the ratio of the top floor displacement RMS in the controlled structure to the top floor displacement RMS in the structure equipped with an optimally tuned passive TMD. Values smaller than 1.0 denote a better effectiveness in comparison to the passive system. To account for the random

character of the earthquake-like base excitation, 1000 episodes of 2000 time steps are simulated for each noise level. Fig. 3 plots the mean value of the RMS ratio (blue line) together with its 1 sigma band (yellow).
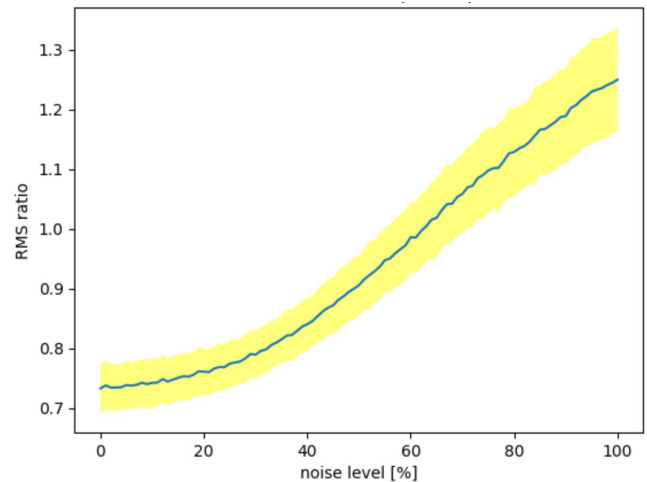

Fig 3. Control efficiency for various measurement noise levels, assessed in terms of the ratio of the top floor displacement RMS between the controlled structure and the structure equipped with an optimally tuned passive TMD: mean value (blue) and 1 sigma band (yellow)

The next evaluation assesses how the trained agent handles model errors, which are possible deviations of physical engineering structures from their ideal mathematical models. The stiffness and mass of individual floors are subject to random change. The generated error follows a normal distribution, limited at 10% of the original level to avoid near-zero or negative values. The evaluation results are presented in Fig. 4. Similarly as in Fig 3, the figure depicts the top floor displacement mean RMS ratio (blue) and its 1 sigma range (yellow), evaluated at each error level using 1000 episodes of 2000 steps each.
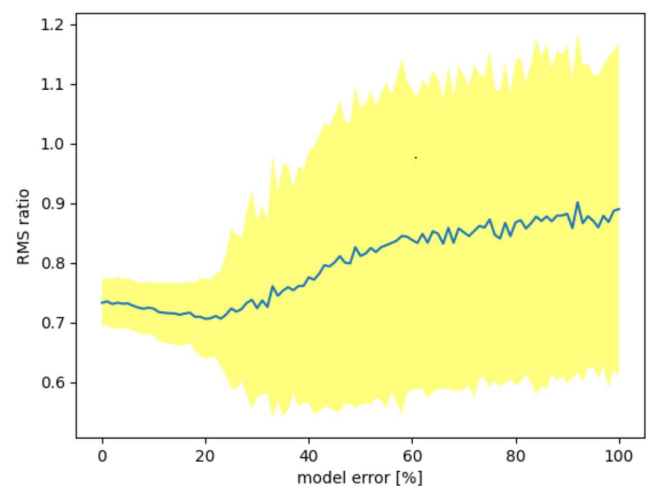

Fig 4. Control efficiency for various model error levels, assessed in terms of the ratio of the top floor displacement RMS between the controlled structure and the structure equipped with an optimally tuned passive TMD: mean value (blue) and 1 sigma band (yellow)

The control performance was not significantly affected by even large levels of measurement noise and model error, as indicated by the observed stable RMS ratio. In particular, the control was more effective than the optimal passive TMD up to measurement noise of about 60% rms. In case of model errors, the mean control effectiveness remained surprisingly good in the entire tested error range; however, the variability of the results increased considerably for model errors above the level of 20%. Such results suggest that the trained model possesses a certain degree of tolerance to disturbances in the form of measurement noise and model errors, allowing it to maintain reliable performance even in the presence of real-world environmental variations.

One possible reason for the model's low sensitivity to disturbances could be attributed to its neural network architecture. Neural networks, particularly those with deeper structures, are known for their ability to learn and extract meaningful features from noisy data. The network layers and parameters might have been optimized during training to capture relevant patterns and generalize well, enabling the model to disregard irrelevant noise components. Additionally, the random character of the base excitation could also contribute to the model resilience, as it prevents the agent from overfitting the specific characteristics of the model and signals and allows it to explore the entire control space.

Further analysis and experimentation can provide deeper insights into the model's robustness and shed light on the specific architectural and training aspects that contribute to its noise tolerance. Understanding these factors will not only enhance our understanding of the model's behavior but also guide the development of more resilient and reliable models in various scientific and engineering domains.

VI. Conclusion

This contribution studied the efficiency of an RL-based semi-active control scheme applied to a shear-type building subjected to an earthquake-like base excitation. The evaluation revealed a noteworthy characteristic of the control system, namely its remarkable insensitivity to measurement and model errors. Despite potential deviations or inaccuracies in the mathematical model used for control, the system demonstrated a high level of robustness and stability. This implies that the control algorithm could effectively compensate for discrepancies between the actual system behavior and the idealized mathematical representation, ensuring reliable performance in real-world scenarios. The observed low insensitivity to noises and errors highlights the effectiveness and practical applicability of the RL-based control methodology in the considered context.

The promising results provide initial insights into the potential of reinforcement learning for improving and ensuring the performance of semi-active damping systems, even though the use of RL in structural control, particularly in semi-active control, is not yet widespread.

References

[1] B. F. Spencer Jr., S. Nagarajaiah. 2003. "State of the art of structural control." *J. Struct. Eng.* 129:845–856, https://doi.org/10.1061/(ASCE)0733-9445(2003)129:7(845)

[2] B. Basu, O. S. Bursi, F. Cascati, S. Cascati, A. E. Del Grosso, M. Domaneshi, L. Faravelli, J. Holnicki-Szulc, H. Irshik, M. Krommer, M. Lepidi, A. Martelli, B. Ozturk, F. Pozo, G. Pujol, Z. Rakicevic, and J. Rodellar. 2014. "A European Association for the Control of Structures joint perspective. Recent studies in civil structural control across Europe," *Struct. Contrl Health Monit.* 21:1414–1436, https://doi.org/10.1002 / stc.1652

[3] F. Casciati, J. Rodellar, and U. Yildirim. 2012. "Active and semi-active control of structures-theory and applications: A review of recent advances." *J. Intell. Mater. Syst. Struct.* 23:1181–1195, https://doi.org/10.1177/1045389X12445029

[4] N. R. Fisco, H. Adeli. 2011. "Smart structures: Part I—Active and semi-active control." *Sci. Iran.* 18(3):275–284, https://doi.org/10.1016/j.scient.2011.05.034

[5] M. Gutierrez Soto, H. Adeli. 2013. "Tuned mass dampers." *Arch. Comput. Meth. E.* 20:419–431, https://doi.org/10.1007/s11831-013-9091-7

[6] S. Elias, V. Matsagar. 2017. "Research developments in vibration control of structures using passive tuned mass dampers." *Annu. Rev. Control* 44:129–156, https://doi.org/10.1016/j.arcontrol.2017.09.015

[7] S. Pourzeynali, H. H. Lavasani, and A. H. Modarayi. 2007. "Active control of high rise building structures using fuzzy logic and genetic algorithms." *Eng. Struct.* 29:346–357, https://doi.org/10.1016/j.engstruct.2006.04.015

[8] D. E. Kirk. *Optimal control theory.* Courier Corporation, 2004.

[9] G. Rypeść, Ł. Lepak, P. Wawrzyński. 2022. "Reinforcement Learning for on-line Sequence Transformation," in *Proc. 17th Conf. on Computer Science and Intelligence Systems*, Sofia, ACSIS 30, pp. 133–139. https://doi.org/10.15439/2022F70

[10] D. Silver, T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, T. Lillicrap, K. Simonyan, and D. Hassabis. 2018. "A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play." *Science* 362(6419):1140–1144, https://doi.org/10.1126/science.aar6404

[11] A. El Sallab, M. Abdou, E. Perot, and S. Yogamani. 2017." Deep reinforcement learning framework for autonomous driving." in *Proc. IS&T International Symposium on Electronic Imaging Science and Technology*, Burlingame, 2017, pp. 70–76, https://doi.org/10.2352/ISSN.2470-1173.2017.19.AVM-023

[12] G. Reddy, A. Celani, T. Sejnowski, and M. Vergassola. 2016." Learning to soar in turbulent environments." *PNAS* 113(33):E4877–E4884, https://doi.org/10.1073/pnas.1606075113

[13] H. Shi, Y. Zhou, X. Wang, S. Fu, S. Gong, and B. Ran. 2022. "A deep reinforcement learning-based distributed connected automated vehicle control under communication failure." *Comput.-Aided Civ. Inf.* 37(15):2033–2051, https://doi.org/10.1111/mice.12825

[14] A. Bernard, I. F. Smith. 2008. "Reinforcement learning for structural control." *J. Comput. Civil Eng.* 22(2):133–139, https://doi.org/10.1061/(ASCE)0887-3801(2008)22:2(133)

[15] A. Khalatbarisoltani, M. Soleymani, and M. Khodadadi. 2019. "Online control of an active seismic system via reinforcement learning." *Struct. Contrl Health Monit.* 26(3):e2298, https://doi.org/10.1002/stc.2298

[16] Z.-C. Qiu, G.-H. Chen, and X.-M. Zhang. 2021. "Reinforcement learning vibration control for a flexible hinged plate." *Aerosp. Sci. Technol.* 118:107056, https://doi.org/10.1016/j.ast.2021.107056

[17] F. L. Lewis, D. Vrabie, and K. G. Vamvoudakis. 2012. "Reinforcement learning and feedback control: Using natural decision methods to design optimal adaptive controllers." *IEEE Contr. Syst. Mag.* 32(6):76–105, https://doi.org/10.1109/MCS.2012.2214134