Journée Régionale des Doctorants en Automatique, 24 juin 2025

Actor-Critic learning for variable damping injection

Alejandro TEVERA RUIZ[1, 2,*], Sergio URZÚA CORREA[2], Vicente PARRA VEGA[2],
Anand SÁNCHEZ ORTA[1], Rodolfo GARCIA RODRÍGUEZ[3], Pedro CASTILLO[1]

[1] *Université de technologie de Compiègne, CNRS, Heudiasyc, 60200 Compiégne, France.*
[2] *Cinvestav Unidad Saltillo, 25900 Ramos Arizpe, Mexique.*
[3] *Universidad Autónoma de Coahuila, 25280 Saltillo, Mexique.*

* *ateverar@hds.utc.fr*

Tuning feedback gains constitute a significant issue for nonlinear systems, particularly for the large class of Lagrangian systems, because it represents a wide variety of physical systems across disciplines. The fundamental difficulty of computing the precise values of these gains stems from the fact that typically exact knowledge of system dynamics is required [1]; given that in practice such knowledge is unavailable, conservative upper bounds are set, used in compliance with Lyapunov stability, [2]; however, it yields larger than needed damping at the expense of increasing the sensibility of the system by increasing its natural frequency.

In particular, quadrotors have been ubiquitously applied to various personal and professional activities related to human daily living because of their flying capabilities; however, there remains a subject of intensive basic and applied research on how to tune their controllers efficiently. Typically, they are tuned with fixed gains, including the damping gain, for a given initial condition. This leads to higher than required gains, increasing sensitivity, and exacerbating energy consumption. Then, the question arises of how to design a variable damping gain to yield the expected performance without higher gains, consequently without unnecessarily higher battery demand.

In this work [3], we explore the Actor-Critic (AC) scheme considering the timely evaluation of error manifold performance to tune damping gains for an chattering-less sliding mode controller for quadrotor’s attitude, [4]. Without any pre-training of the Neural Networks (NN) neither of adaptive parameters nor any knowledge of dynamics, local tracking is guaranteed with robustness against disturbances.

The idea behind the AC scheme is to design a regime of injecting higher damping when far from the target but lower damping otherwise, up to a lower bound. In addition, if a convenient damping value has already been computed, but a sudden unknown disturbance affects the system, the AC motor learning mechanism must react automatically, decreasing or increasing the damping according to the reward policy until the disturbance disappears. In this way, we design a novel AC scheme including a new idea of ``virtual control" to enforce better learning, for the Critic NN, where the value function is approximated using a policy reward provided by an expert user to satisfy the sliding mode condition. Thus, weights of the Critic NN are designed to be driven by a non-linear saturated monitor, which oversees how far the value function approximation variation is from the target. This monitor is vital for instrumenting the required policy to yield smooth tuning of damping injection by means of the Actor NN.

The numerical simulations demonstrate that the proposed scheme significantly reduces energy consumption compared to the fixed damping strategy. This reduction can be attributed to two primary factors. Firstly, our scheme applies less initial power than the fixed one due to the lower weight tracking error, resulting in a smoother increase in damping and power. Second, while the fixed damping approach applies the same energy throughout, sometimes even more than necessary, the variable damping strategy adapts to the sliding mode condition, modifying the applied power only as needed. This approach leads to a decrease in energy consumption and improves the overall robustness, even under the influence of substantial disturbances.

Therefore, the net result is a less conservative tuning approach, which still depends on the Lyapunov stability. Our ongoing work focuses on developing neuro-control strategies with theoretical guarantees and experimental validation to handle aerodynamic disturbances during quadrotor navigation in complex environments.

Références :

[1] Borase, R.P. and Maghade, D.K. and Sondkar, S.Y. and others, A review of PID control, tuning methods and applications. *International Journal of Dynamics and Control*, 2021, 9:818--827.
[2] Kelly, R., A tuning procedure for stable PID control of robot manipulators. *Robotica*, 1995, 13(2):141--148.
[3] Tevera-Ruiz, A. and Urzúa-Correa, S. E. and Parra-Vega, V. and Sanchez-Orta, A. and Garcia-Rodriguez R., "Actor-Critic Learning of Variable Damping Injection for Quadrotor Attitude Robust Control," 2023 20th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE), Mexico City, Mexico, 2023, pp. 1-6, doi: 10.1109/CCE60043.2023.10332878.
[4] Sanchez-Orta, A. and Parra-Vega, V. and Izaguirre-Espinosa, C. and Garcia, O., Position--Yaw Tracking of Quadrotors. *Journal of Dynamic Systems, Measurement, and Control*, 2015, 137(6):061011.