



HAL
open science

Exploration strategies for control of chaotic dynamical systems using reinforcement learning

Rémy Hosseinkhan Boucher, Amine Saibi, Michele Alessandro Bucci, Onofrio Semeraro, Lionel Mathelin

► **To cite this version:**

Rémy Hosseinkhan Boucher, Amine Saibi, Michele Alessandro Bucci, Onofrio Semeraro, Lionel Mathelin. Exploration strategies for control of chaotic dynamical systems using reinforcement learning. European Drag Reduction and Flow Control Meeting - EDRFCM 2022, Sep 2022, Paris, France. hal-03854602

HAL Id: hal-03854602

<https://hal.science/hal-03854602v1>

Submitted on 15 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

EXPLORATION STRATEGIES FOR CONTROL OF CHAOTIC DYNAMICAL SYSTEMS USING REINFORCEMENT LEARNING

Rémy Hosseinkhan Boucher

LISN, CNRS, University Paris-Saclay, 91400 Orsay, France

Amine Saibi

JR-D’Alembert, UPMC–Sorbonne Universit, 75005 Paris, France
LISN, CNRS, University Paris-Saclay, 91400 Orsay, France

Michele Alessandro Bucci

TAUteam, INRIA Saclay, LISN, Université Paris-Saclay, 91400 Orsay, France

Onofrio Semeraro

LISN, CNRS, University Paris-Saclay, 91400 Orsay, France

Lionel Mathelin

LISN, CNRS, University Paris-Saclay, 91400 Orsay, France

Numerous research efforts have been devoted to the application of control theory to fluid flows in the last decades. Common approaches consist of reduced order models of the original system, often within the linear limit, for enabling the control design based on standard control theory. However, these approximated representations are often the main limitation of standard applications, as these control techniques usually exhibit poor performance when applied in off-control design conditions; an example is provided by the impact that non-linearities have on the dynamics of the closed-loop when neglected [9].

REINFORCEMENT LEARNING

An alternative is represented by fully data-driven methods where a physical model is not employed; Reinforcement Learning (RL) algorithms allow such a strategy while preserving optimality of the control solutions. Indeed, some RL algorithms can be regarded as a fully data-driven counterpart of the discrete-in-time optimal control strategies based on the Bellman equation, where the policy – i.e., the control action – is directly learnt from interactions of the system with the environment without relying on an a-priori physical model [10]. The relevance of a control action when the system is in a given state is measured by a reward and the expected cumulative rewards (or Value function) is the objective to be maximized. When neural networks are employed as an approximation format, the framework is referred to as deep RL (DRL). Recent works show the feasibility of the approach for the control of instabilities classically found in fluid mechanics systems [7, 2, 6].

IMPROVING LEARNING BY IMPROVING THE EXPLORATION

The methods we consider here are fully data-driven, without prior knowledge of a model. This implies that the modelling step is now replaced by the so-called *exploration*. Since the dimension of the phase space of the system under consideration is very large with respect to traditional tasks in

Reinforcement Learning, the data efficiency of state-of-the-art methods such as Proximal Policy Optimization [8] and Deep Deterministic Policy Gradient [5] must be improved.

As an illustration in this work, we will consider the system governed by the Kuramoto-Sivashinsky (KS) equations. This system is described by a partial differential equation with initial value and spatial periodicity conditions:

$$\frac{\partial u}{\partial t}(t, x) = -u(t, x) \frac{\partial u}{\partial x}(t, x) - \frac{\partial^2 u}{\partial x^2}(t, x) - \frac{\partial^4 u}{\partial x^4}(t, x), \quad (1)$$

with an initial condition $u(0, x)$, and periodicity in space $u(t, x + M) = u(t, x)$ given a length of the domain $M \in \mathbb{R}_+$. This 1-D dynamical system exhibits a chaotic behavior for a range of values of M and is often used as a convenient proxy of turbulent flow regimes. As such, it has recently been used as a practical tool to evaluate the robustness of machine learning methods. Properties of the KS system as discussed in detail in [4]. As an illustration, a controllable trajectory in the chaotic regime is depicted in Fig. 1. In a previous work [3], we have shown that using the Deep Deterministic Policy Gradient (DDPG) algorithm allows to stabilize these chaotic trajectories; Fig. 2 shows an illustration of the method. More precisely, the controller task is to drive the trajectory of the chosen dynamical system to transit from one system equilibrium to another one by accumulating statistical knowledge of the solution space (environment). The resulting policies were however far from being the optimal ones discussed in [1]. In this contribution, we hence intend to understand and identify the multiple factors and their influence on the policy identification, with a quantitative scope, but without relying on physics constraints or prior knowledge. In particular, we develop consistent exploration strategies which allow to carefully sample new observations, based on the previous ones, in order to achieve the knowledge of the system to be controlled. As a result, better control policies – in principle, the optimal ones – could be obtained with a lower amount of interactions with the system. From this perspective, we rely on *information theory* and *optimal design of experiments* to quantify the knowledge the controller gathers during the process.

We will compare the control policies achieved with the different strategies, and contrast them with the standard linear solutions. More generally, the final goal is to improve the use of data-driven and statistical learning for computational fluid dynamics, and specifically reinforcement learning. This approach, at the crossing between machine learning and dynamical systems, can also be of help to the streaming learning community to gather insights for explaining and understanding behaviour of data-based modelling.

REFERENCES

[1] A. Armaou and P. D. Christofides. Feedback control of the kuramoto–sivashinsky equation. *Physica D: Nonlinear Phenomena*, 137(1-2):49–61, 2000.

[2] G. Beintema, A. Corbetta, L. Biferale, and F. Toschi. Controlling rayleigh–bénard convection via reinforcement learning. *Journal of Turbulence*, 21(9-10):585–605, 2020.

[3] M. A. Bucci, O. Semeraro, A. Allauzen, G. Wisniewski, L. Cordier, and L. Mathelin. Control of chaotic systems by deep reinforcement learning. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, pages 278–285, 2019.

[4] P. Cvitanović, R. L. Davidchack, and E. Siminos. On the state space geometry of the kuramotosivashinsky flow in a periodic domain. *SIAM Journal on Applied Dynamical Systems*, 9:1–33, 2010.

[5] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra. Continuous control with deep reinforcement learning. 2015.

[6] R. Paris, S. Beneddine, and J. Dandois. Robust flow control and optimal sensor placement using deep reinforcement learning. *Journal of Fluid Mechanics*, 913, 2021.

[7] J. Rabault, M. Kuchta, A. Jensen, U. Rglade, and N. Cerardi. Artificial neural networks trained through deep reinforcement learning discover control strategies for active flow control. *Journal of Fluid Mechanics*, 865:281302, 2019.

[8] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov. Proximal policy optimization algorithms. 2017.

[9] D. Sipp and P. J. Schmid. Linear Closed-Loop Control of Fluid Instabilities and Noise-Induced Perturbations: A Review of Approaches and Tools. *Applied Mechanics Reviews*, 68, 2016.

[10] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. The MIT Press, 2018.

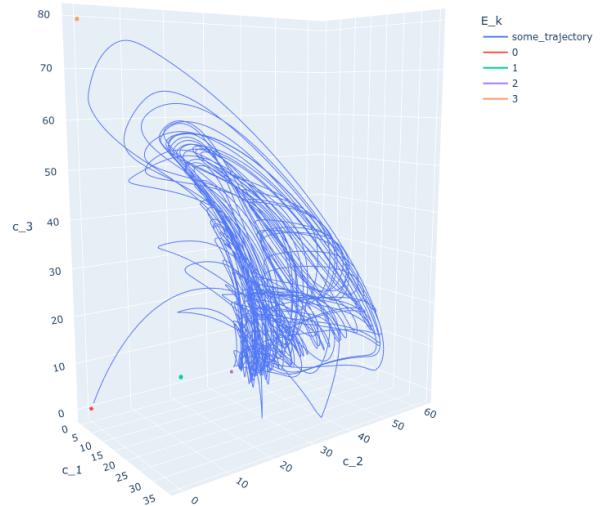


Figure 1: A sampled chaotic trajectory emanating from the neighbourhood of the equilibrium $u = 0$ of the Kuramoto-Sivashinsky equation under chaotic regime ($M = 22$); the visualization is obtained by projecting on the first 3 Fourier modes. The 3 non-trivial equilibria (stationary solutions) are also shown.

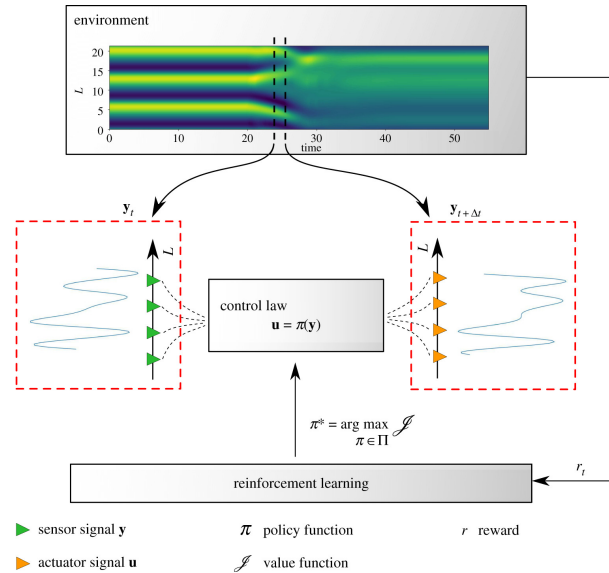


Figure 2: Reinforcement Learning scheme: from an initial state, the control law applies a forcing to the state of the dynamical system. Given this forcing action (or control) and the new state, a reward is obtained. An optimal reinforcement learning policy is a set of control decisions which maximises the asymptotical sum of rewards. In this case, a reward or cost could be the energy of the system. Taken from [3].