Deep reinforcement learning control of flow over rotary oscillating cylinder at low Reynolds number

Mikhail P. Tokarev
Novosibirsk State University and
Institute of Thermophysics SB RAS
Novosibirsk, Russia
mtokarev@itp.nsc.ru

Evgeniy N. Pavlovskiy
Novosibirsk State University
Novosibirsk, Russia
e.pavlovskiy@g.nsu.ru

Egor V. Palkin
Novosibirsk State University and
Institute of Thermophysics SB RAS
Novosibirsk, Russia
palkinev89@gmail.com

Rustam I. Mullyadzhanov
Novosibirsk State University and
Institute of Thermophysics SB RAS
Novosibirsk, Russia
r.mullyadzhanov@g.nsu.ru

Abstract

We apply deep reinforcement learning to active closed-loop control of a two-dimensional flow over a cylinder oscillating around its axis with a time-dependent angular velocity representing the only control parameter. Experimenting with the angular velocity, the neural network is able to devise a control strategy based on low frequency harmonic oscillations with some additional modulations to stabilize the Kármán vortex street at low Reynolds number $Re = 100$ leading to the drag reduction of 14%. The additional efforts are very low as the maximum amplitude of the angular velocity does not exceed 10% of the incoming flow. A detailed comparison with a flow controlled by harmonic oscillations with one frequency and a fixed amplitude are presented highlighting the necessity of a feedback loop.

1 Introduction

A well-known Kármán vortex street is typically formed in the wake of the flow over a bluff body exerting an oscillating value of the force [1]. This unsteadiness may cause structural damages due to the coupling of the body vibrations and pressure fluctuations of the fluid. Over time a large number of flow control strategies are proposed to influence and suppress these unwanted dynamical features [2]. The overall classification includes passive and active methods due to the possible energetic input to the flow [3]. Active methods represent open- and closed-loop control depending on the presence of a feedback from sensors to actuators with a further update of a control signal. An appealing way to design new closed-loop control strategies is to rely on the so-called data-driven and learning-based methods which lately receive well-deserved attention [4].

In fluid dynamics machine learning techniques have been fruitfully applied to the issue of the turbulence closure modelling within Large-eddy simulations and Reynolds-averaged Navier–Stokes equations, reconstructing flow fields, recovering dynamical features and control [4]. A promising way to devise new control strategies is a combination of multilayer (deep) neural networks combined with the reinforcement learning (RL) strategy [5] resulting in a deep reinforcement learning (DRL) paradigm succeeding in a large number of problems [6]. RL represents a self-learning strategy introducing an agent who interacts with the environment through particular actions in order to get a maximum reward. Recent examples consider the manipulation of the flow over a cylinder

Third Workshop on Machine Learning and the Physical Sciences (NeurIPS 2020), Vancouver, Canada.

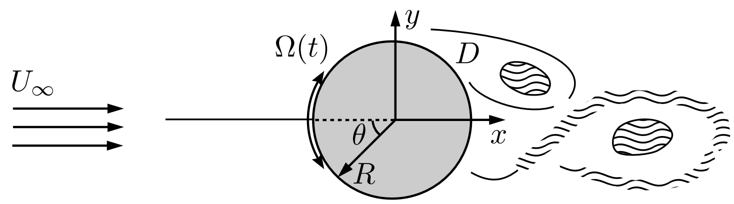

Figure 1: A rotary oscillating cylinder in a cross-flow with a coordinate system.

using multiple synthetic jets and numerical simulations at low Reynolds numbers delivering robust DRL-based control strategies [7-8].

In this work we study a closed-loop control strategy for the flow over a cylinder rotating around its axis with the time-dependent angular velocity being the only control parameter while in the mentioned papers [7-8] flow control was implemented using synthetic jets on a fixed cylinder. We employ the direct numerical simulations (DNS) of the Navier–Stokes equations at low Reynolds number $Re = 100$ and apply the DRL method which relies on the Proximal policy optimization algorithm [9] maximizing the expected reward. The reward value is based on the lift and drag forces acting on the cylinder with the neural network employing the information on the pressure field in the wake region. The particular focus of the work is on the deviation of the control signal from the intuitive one typically considered as harmonic oscillations.

2 Problem formulation and computational details

We consider a cylinder of the diameter D in a fluid cross-flow with a uniform incoming velocity U_∞, see Fig. 1. The considered Reynolds number $Re = U_\infty D/\nu = 100$ represents the flow regime with a Kármán vortex street behind the bluff body, where ν is the kinematic viscosity. The applied control strategy is based on the rotation of the cylinder around its axis with the wall velocity $U_w(t) = U_\infty\Omega(t)$. The primary goal is to find the optimal signal $\Omega(t)$ to influence drag and lift coefficients:

$$C_D = \frac{2F_x}{\rho U_\infty^2}, \qquad C_L = \frac{2F_y}{\rho U_\infty^2}, \tag{1}$$

where F_x and F_y are the drag and lift forces acting on the cylinder per unit length and ρ is the fluid density.

2.1 Flow computations

To describe the flow field we solve the non-dimensional incompressible Navier–Stokes equations:

$$\frac{\partial u_i}{\partial t} + u_j\frac{\partial u_i}{\partial x_j} = -\frac{\partial p}{\partial x_i} + \frac{1}{Re}\frac{\partial^2 u_i}{\partial x_j^2}, \qquad \frac{\partial u_j}{\partial x_j} = 0, \tag{2}$$

where u_i and p are the velocity components and pressure field and all quantities are non-dimensionalized using U_∞ and D. The configuration is considered in a two-dimensional setup with the coordinate system located in the center of the cylinder and x, y representing the streamwise and vertical coordinates, respectively. The computations are performed using an open-source unstructured finite-volume code T-Flows [10] below referred to as the CFD solver. The computational domain represents a rectangle of the size $L_x \times L_y = 30D \times 20D$ with the center of the cylinder placed $10D$ from the inlet boundary. The slip condition is set at the top and bottom boundaries and convective outflow condition is prescribed at the outlet, while control law representing the tangential velocity of the cylinder wall is described below. The mesh contains 15140 hexahedral cells while the computational timestep is $\Delta t_{CFD} = 10^{-2}$. We validate the results of the stationary cylinder case. The typical quantities of interest are the time-averaged drag coefficient $\overline{C}_D = 1.33$ and vortex shedding frequency $f_{vs} = 0.17$ which are in excellent agreement with the data from the literature [1].

2.2 Machine-learning architecture and feedback loop

Below we describe a set of algorithms employed to obtain the angular velocity signal $\Omega(t)$ with the closed-loop controller synthesized by training a fully connected neural network (FNN). The optimal

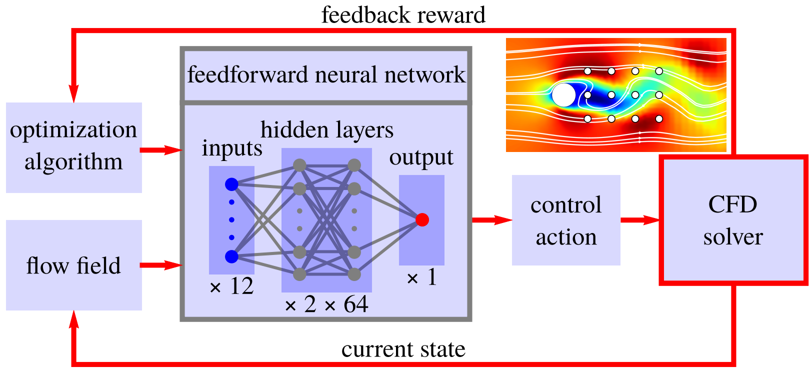

Figure 2: Active closed-loop flow control optimization scheme.

control strategy evaluation relies on the Reinforcement learning (RL) approach [5] and a Policy gradient (PG) algorithm [9] with a maximization of a defined reward function. A schematic view of optimization of the feedback control system is shown in Figure 2. In RL the environment representing the CFD solver interacts with the agent which is the FNN controller in our case. The agent takes a new action based on the current state of the environment representing the data from 4×3 array of pressure sensors placed in the near wake behide the cylinder, see 'inputs' and the flow schematics in Figure 2. The FNN has two hidden layers with 64 neurons each and a single output for the mean of the policy PDF of the cylinder angular velocity during evaluation of the controller. However, training employs two networks of the similar structure for policy and value prediction including a trainable dispertion coefficient of the policy PDF for a smooth transition from exploration to exloitation of the learnt policy. The number of neurons was determined experimentally obseving the learning speed and the mean reward after convergence with a fixed number of inputs and the reward function. The neurons in the hidden layers used the sigmoidal activation function while the activation of the output neuron was linear. The action defined the next value of Ω which was put forward to the CFD solver with a specific relaxation procedure in time. The action time step T_{ac} was set small enough compared to the characteristic time scale $T_{vs} = 1/f_{vs}$ of the vortex shedding according to recommendations [7]. We employed the value $T_{ac} \approx 0.05T_{vs}$ with $T_{ac} = 30\Delta t_{CFD} = 0.3$. After receiving a new target value of Ω, the angular velocity of the cylinder was updated from the old value to a new target one linearly in time within the whole interval T_{ac}. We performed additional tests with the exponential relaxation scheme [7] as well as step-like change of Ω, however, a better behaviour of the linear one for a test harmonic control law was observed. Typical learning process required full four days with over a hundred thousand action steps made in parallel using three vectorized environments locally on a PC with 24 cores in total. The convergence was reached after three days of training with near eight thousand control steps. As a parameter for the control optimization algorithm we chose the learning rate constant of the value 3×10^{-4}, clip parameter $\epsilon = 0.2$, GAE parameters with discount rate $\gamma = 0.99$ and $\lambda = 0.95$, value coefficient $c_1 = 0.5$, entropy coefficient $c_2 = 0.01$ [9]. The instantaneous reward value was parametrized in the following form:

$$r = R_1 - (\langle C_D \rangle_{ac} + R_2 |\langle C_L \rangle_{ac}|), \tag{3}$$

where the drag and lift coefficients were averaged over the action time step. For convenience the constant R_1 was chosen so that the reward values per action are close to zero. The non-zero constant R_2 constrained the non-zero time-averaged lift leading to the asymmetric flow regimes. With $R_2 = 0$ in the long run the control law tends to return a rotating regime. We chose the final values as $R_1 = 3$ and $R_2 = 0.1$ [7].

3 Results and discussion

The training process required at least 80 epochs, i.e. the intervals between control actions resulting in the overall time interval of near 35000 time units in terms of D/U_∞. The evolution of the reward value averaged over the action time step $\langle r \rangle_{ac}$ reached the saturation after around 50 epochs while the entropy function decreased monotonically. After the training process the performance of the neural controller was evaluated. Figure 3 shows the instantaneous streamwise velocity field around the cylinder for the case without control and after 80 time units we applied the DRL-based scheme which is enough to reach a new steady state. As a result the flow stabilizes with a small rotational input trying to balance this inherent instability. The length of the recirculation bubble becomes larger with the local pressure minimum moving further downstream. This leads to a smaller pressure difference between the front and rear of the cylinder with a final decrease of C_D.

Figure 4 shows the evolution of several characteristics such as the angle of a certain point on the surface of the cylinder θ, angular velocity Ω, drag and lift coefficients C_D and C_L for different flow regimes. The case of a stationary cylinder is demonstrated as a reference one with $\theta = \Omega = 0$. The drag coefficient exhibits sinusoidal behaviour with a frequency $2f_0$ around a time-averaged value $\overline{C}_D = 1.33$ while C_L fluctuates with a natural frequency $f_0 = 0.17$ reflecting the vortex shedding process as mentioned above. The application of the DRL-scheme for control leads to a transient period of about 20 time units resulting in modulated harmonic signal of the angular velocity Ω. However, the angle θ also evolves sinusoidally in time with a linear trend of a very small slope corresponding to a negligible rotation angular velocity in one direction. Note that the Fourier spectrum of the C_L signal shows the main peak at a frequency $f \approx 0.6 f_0$ with small secondary peaks at higher and lower frequencies giving room for modulations of Ω. These modulations turn out to be essential for a drag reduction which is demonstrated in a straightforward manner. We apply a harmonic forcing to the cylinder with $\Omega = \Omega_0 \sin(2\pi f t)$ where $f = 0.6 f_0$, $\Omega_0 = 0.08$ well approximates the signal from a DRL-scheme. The outcome of this control strategy is a slight decrease of $\overline{C}_D = 1.29$ compared to the stationary case and is far from a DRL-scheme $\overline{C}_D = 1.15$ (-14% from the case with the fixed cylinder). Thus, the feedback loop is indeed gives a benefit for active flow control correctly responding on the instantaneous phase of the vortex shedding process by tuning the angular velocity of the cylinder to stabilize the wake and decrease the drag.

4 Conclusion

We applied deep reinforcement learning to active closed-loop control of a two-dimensional flow over a cylinder oscillating around its axis with a time-dependent angular velocity representing the only control parameter. Probing different values of the angular velocity, the neural network was able to create a control strategy based on low frequency harmonic oscillations with some additional modulations to stabilize the Kármán vortex street at low Reynolds number $Re = 100$ leading to the drag reduction of 14% comparable with a state-of-the-art control theory optimization routines based on adjoint methods. The additional input of energy to rotate the cylinder was very low as the maximum amplitude of the angular velocity did not exceed 10% of the incoming flow. A detailed comparison with a flow controlled by harmonic oscillations with one frequency and a fixed amplitude were presented highlighting the necessity of a feedback loop.

Broader Impact

The results of this research may be interesting for development of new transportation and energy systems with an intelligent active flow control in case the approach can be adapted for realistic flow conditions, i.e. turbulent flows with higher Reynolds numbers, and significant nonlinear effects is of

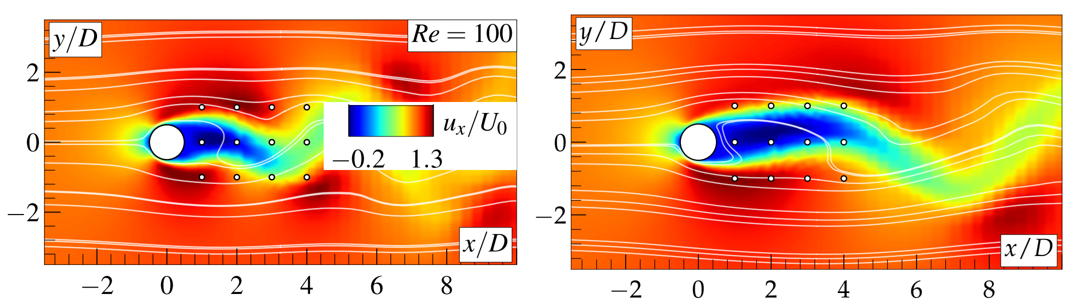

Figure 3: Typical instantaneous streamwise velocity field with streamlines. Left: stationary cylinder. Right: DRL-based flow regime after 80 time units of control. The array of 4×3 white points corresponds to pressure sensors serving as the input for the neural network, see Figure 2.

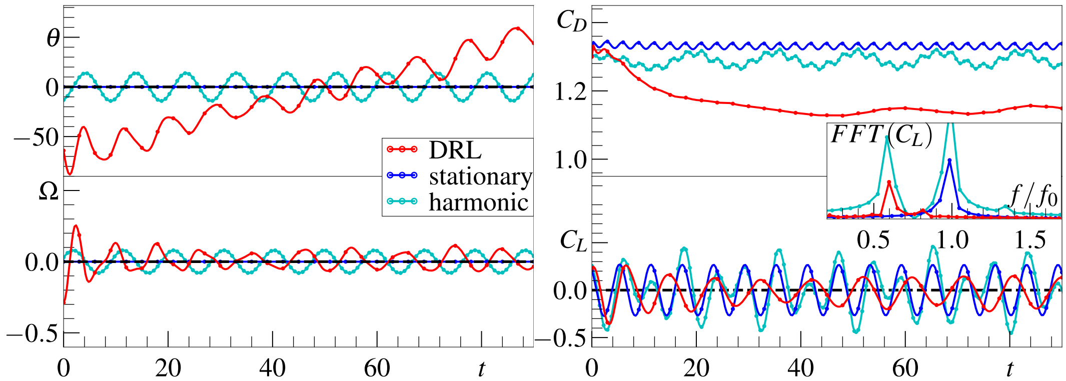

Figure 4: Rotation angle and angular velocity (left) with the drag and lift coefficients (right). Three flow regimes are shown, i.e. the stationary cylinder (blue line), with a DRL-scheme for control (red), forced with harmonic oscillations with $\Omega(t) = \Omega_0 \sin(2\pi f t)$ where $\Omega_0 = 0.08$, $f = 0.6 f_0$ (green line).

particular interest as it is difficult to account by the conventional linear control over a wide range of parameters. Currently, direct neural controllers are almost not used in industrial systems, since their stability and reliability have not been proven. This can lead to a possible accident if used without adequate accident prevention systems.

Acknowledgments and Disclosure of Funding

The work is partly supported by Russian Science Foundation grant No. 19-79-30075 (EP and RM for CFD simulations) and Russian Foundation for Basic Research grant No. 20-08-01093 (MT for coupling of the machine learning optimization algorithm to active flow control with a feedback loop). We also acknowledge the support of the Mathematical Center in Novosibirsk State University and a pleasant environment.

References

[1] Williamson, C.H. (1996) Vortex dynamics in the cylinder wake. *Annual review of fluid mechanics* **28**(1):477-539.

[2] Choi, H., Jeon, W.P., Kim, J. (2008) Control of flow over a bluff body. *Annual review of fluid mechanics* **40**:113–139.

[3] Gad-el Hak, M. (2007) Flow control: passive, active, and reactive flow management: Cambridge university press

[4] Brunton, S.L., Noack, B.R. & Koumoutsakos, P. (2020) Machine learning for fluid mechanics. *Annual Review of Fluid Mechanics* **52**:477-539.

[5] Sutton, R.S. & Barto, A.G. (2018) Reinforcement learning: An introduction. MIT Press.

[6] 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. (2015) Human-level control through deep reinforcement learning. *nature* **518**(7540):529-533.

[7] Rabault, J., Kuchta, M., Jensen, A., Réglade, U. & Cerardi, N. (2019) Artificial neural networks trained through deep reinforcement learning discover control strategies for active flow control. *Journal of Fluid Mechanics* **865**:281-302.

[8] Rabault, J. & Kuhnle, A. (2019) Accelerating deep reinforcement learning strategies of flow control through a multi-environment approach. *Physics of Fluids* **31**(9):094.

[9] Schulman, J., Wolski, F., Dhariwal, P., Radford, A. & Klimov, O. (2017) Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*

[10] Ničeno, B., Palkin, E., Mullyadzhanov, R., Hadžiabdić, M. & Hanjalić, K. (2018) *T-Flows Web page* https://github.com/DelNov/T-Flows