

From active learning to deep reinforcement learning: Intelligent active flow control in suppressing vortex-induced vibration

Cite as: Phys. Fluids **33**, 063607 (2021); <https://doi.org/10.1063/5.0052524>

Submitted: 31 March 2021 . Accepted: 24 May 2021 . Published Online: 16 June 2021

Changdong Zheng (郑畅东), Tingwei Ji (季廷炜), Fangfang Xie (谢芳芳), Xinshuai Zhang (张鑫帅), Hongyu Zheng (郑鸿宇), and Yao Zheng (郑耀)

COLLECTIONS

Paper published as part of the special topic on [Selected Papers from the 11th National Congress on Fluid Mechanics of China](#)



[View Online](#)



[Export Citation](#)



[CrossMark](#)

ARTICLES YOU MAY BE INTERESTED IN

[Applying deep reinforcement learning to active flow control in weakly turbulent conditions](#)
Physics of Fluids **33**, 037121 (2021); <https://doi.org/10.1063/5.0037371>

[Mechanical rotation at low Reynolds number via reinforcement learning](#)
Physics of Fluids **33**, 062007 (2021); <https://doi.org/10.1063/5.0053563>

[Physics guided machine learning using simplified theories](#)
Physics of Fluids **33**, 011701 (2021); <https://doi.org/10.1063/5.0038929>

Physics of Fluids
SPECIAL TOPIC: Tribute to
Frank M. White on his 88th Anniversary

[SUBMIT TODAY!](#)

From active learning to deep reinforcement learning: Intelligent active flow control in suppressing vortex-induced vibration

Cite as: Phys. Fluids 33, 063607 (2021); doi: 10.1063/5.0052524

Submitted: 31 March 2021 · Accepted: 24 May 2021 ·

Published Online: 16 June 2021



View Online



Export Citation



CrossMark

Changdong Zheng (郑博东), Tingwei Ji (季廷炜), Fangfang Xie (谢芳芳),^{a)} Xinshuai Zhang (张鑫帅), Hongyu Zheng (郑鸿宇), and Yao Zheng (郑耀)

AFFILIATIONS

School of Aeronautics and Astronautics, Zhejiang University, Hangzhou, Zhejiang 310027, China

Note: This paper is part of the special topic, Selected Papers from the 11th National Congress on Fluid Mechanics of China.

^{a)}Author to whom correspondence should be addressed: fangfang_xie@zju.edu.cn

ABSTRACT

In the present work, an efficient active flow control strategy in eliminating vortex-induced vibration of a cylinder at $Re = 100$ has been explored by two machine learning frameworks, from active learning to reinforcement learning. Specifically, an adaptive control scheme by a pair of jets placed on the poles of the cylinder as actuators has been discovered. In the active learning framework, a Gaussian process regression surrogate model is used to predict vibration amplitude of the cylinder using a limited number of numerical simulations by combining the Bayesian optimization algorithm with specified control actions while in the reinforcement learning framework, soft actor-critic deep reinforcement learning algorithm is adopted to construct a real-time control system. The results have shown that the triangle control agent in the active learning framework can reduce the vibration amplitude of the cylinder from $A = 0.6$ to $A = 0.43$. The real-time control in the reinforcement learning framework can successfully suppress the vibration amplitude to 0.11, which is decreased by 82.7%. By comparison, there are some similarities in the amplitude and phase of the action trajectories between two intelligent learning frameworks. They both aim at keeping track of the antiphase between the position and the action, which will restrain the cylinder at a low-amplitude vibration. The underlying physics shows that the jet will contain suction in the stage of vortex generation and injection in the stage of vortex shedding. The current findings have provided a new concept to the typical flow control problem and make it more practical in industrial applications.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0052524>

I. INTRODUCTION

Since Strouhal published his article on bluff bodies¹ in 1878, there have been a large number of experimental studies and theoretical analysis on the vortex structure that appears on the wake of moving objects. In particular, the Newtonian fluid passing through a cylinder gives rise to different interesting phenomena, which is of great concern to scientists. Under a certain flow configuration, the asymmetric vortices shed from the surface of the cylinder generate an unstable lateral force on the cylinder, causing the cylinder to oscillate structurally in turn. That is known as vortex-induced vibration (VIV).²

Vortex-induced vibration (VIV) is often considered in design issues, especially when the vortex shedding frequency is close to the natural frequency of the cylinder, leading to a large-amplitude vibration phenomenon beyond the deformation capacity of the cylinder. Hence, it is necessary to conduct research on the suppression of VIV.^{3,4} From the perspective of the mechanism, there are two effective

ways to perform currently. The first method is to change the strength of the structure to stagger the frequency of vortex shedding, which might deviate from the original design concept. The second method is to change the frequency of vortex shedding by flow control methods so that it staggers the natural frequency of the cylinder to avoid large-amplitude vibration, which is appreciated by most people.

There are two popular flow control methods: the active (active flow control, AFC) and the passive (passive flow control, PFC). The latter is a control method that is fixed at the design stage and cannot be changed during work. The helical strake is a good choice in the field of PFC. With the helical strake on the surface, it can guide the direction of flow and suppress the generation of vortex to a certain extent.⁵ However, the helical strake often brings about an increase in resistance, which means that the rod-shaped structure might encounter dangerous bend along its length. With an additional target of resistance reduction, Galvao *et al.* proposed a foil structure, which

significantly eliminated the vibration and reduced the drag coefficient to about $C_d = 0.5$ under the subcritical Reynolds number.⁶ The free-rotating fairing is one of the current focuses of researchers as well. Whether it is spindle-shaped⁷ or U-shaped fairing,⁸ it is able to control resistance while suppressing vibration. In addition to the above-mentioned methods, scientists have also proposed a passive-nozzle flow control method. The fairing covers over the cylinder surface with holes set at suitable positions where flows can be sucked or injected just like a true nozzle. Chen⁹ and Pontaza¹⁰ have carried out experiments and numerical simulations on this control method.

Considering the disadvantages of PFC, scientists have proposed a more flexible control method, that is, active flow control (AFC). This kind of control mode can be controlled manually and turned on when needed, which has a strong adaptability. Especially for a well-designed wing, adding a fairing component for the purpose of drag reduction will bring about aerodynamic loss. Instead, jet system can avoid such problem easily. In the problem of flow around a cylinder with jet control, Delaunay¹¹ studied the effect of steady base suction and blowing on the stability and dynamics of the cylinder wake. Kim¹² applied a distributed forcing to flow over a circular cylinder for drag reduction by a blowing and suction from the slots located at upper and lower surfaces of the cylinder. However, the complex control device and additional energy input limit the promotion of the AFC. Therefore, compared to the PFC, less work has been done to develop the active control technology, such as, the synthetic jet,¹³ and exploring the active control strategies, e.g., maximum control efficiency. Barbagallo *et al.* established a surrogate model of the flow and used the classic Linear Quadratic Gaussian method (LQG) closed-loop control nozzle system to control the instability of the circular step.¹⁴ The traditional control theory relies on the accuracy of the linearization operation of the nonlinear system, namely, the accuracy of the surrogate model, which is a challenge for the high-dimensional and highly nonlinear complex flow. In recent years, with the advancement of computation and artificial intelligence technology, some intelligent algorithms are introduced to the field of the AFC. As for the step flow problem, Gautier *et al.* used genetic programming to find a reasonable control function to control the actuators with the input collected by the sensors.¹⁵ In order to find an unsupervised learning method close to the optimal control law, Zhou *et al.* use genetic programming method to build a real-time artificial intelligence control system to maximize the mixing rate of turbulent jets.¹⁶ The complexity of the turbulence problem makes it unrealistic to derive the turbulence theory directly, but the development of deep neural networks makes it possible to stride across the turbulence problem by utilizing large datasets. Lee *et al.* used a deep neural network to learn the relationship between wall shear stress and actuator control, which achieved the expected results.¹⁷ As for VIV of cylinder, due to the complexity of the fluid-structure coupling model¹⁸ and the limitation of computation resource, most current researches on AFC focus on the control of constants or harmonics action profile. Dong *et al.* studied the Windward-suction and Leeward-blowing strategy (WSLB) strategy to verify the effectiveness of the control method under different Reynolds numbers (500, 1000), damping coefficients (0.046, 0.0046), and other operating conditions, connected to some PFC methods.¹⁹ Zhu *et al.* replaced the medium sprayed by the nozzle, which has shown that the WSLB strategy using various medium has the effect of suppressing VIV as well.²⁰

Generally, for suppressing VIV, the wind tunnel experiment is a good way to design an effective control strategy and obtain feedback

for every acquisition of the control action space.²¹ On the other hand, with the development of computers and the increase in computational resources, extensive work has been conducted by computational fluid dynamics (CFD),²² while to explore-and-exploit in parametric spaces of high dimensions, it is very time-consuming for traditional approaches of systematic parametric variation²³ in CFD simulations. Therefore, an active learning (AL) framework with Gaussian process regression (GPR) and Bayesian optimization (BO) has been proposed. The active learning framework has used an interrogation function to balance the relationship between exploration and exploitation, and it is expected to solve the global optimization problem with the least computational cost.²⁴ Zheng *et al.* used GPR to establish a surrogate model linking aerodynamic parameters to kinetic parameters of the flapping wing, and the input data mixed with high-resolution and low-resolution CFD results were determined by Bayesian optimization framework.²⁵ Moreover, the active framework has an ability to deal with complex objective functions with noise.²⁴ This ability seems to fit well with the turbulence problem in fluid mechanics, which is computationally expensive and has large uncertainties. Talnikar *et al.* used large eddy simulation method to calculate turbulent channels flow, used BO to minimize turbulence drag, and designed the trailing edge of turbine blades to reduce turbulent heat transfer and pressure loss.²⁶ Mahfoze *et al.* optimized the low-amplitude wall normal blowing control of the turbulent boundary layer flow with an expensive direct numerical simulation method to get a high calculation accuracy.²⁷

Although the AL framework has great advantages in saving computation resources, it is still necessary to manually specify the form of the time-varying control action profile, either a constant or a trigonometric function. The reason why these simple forms are far from meeting the requirements of full exploration is that the complex high-dimensional nonlinear flow state space might correspond to the same complex state space as well. With the rise of artificial intelligence algorithms, especially the emergence of deep reinforcement learning (DRL), it is possible to explore an intelligent strategy by mutually interacting with the environment without knowing any prior knowledge. The DRL framework turns out effectively in Atari's game environment²⁸ and robot operation.²⁹ Therefore, people have introduced DRL as solution to the flow problems. Reddy *et al.* were at the forefront of combining RL framework and fluid mechanics problems. They started from the fact that birds in nature use convectively rising hot air to obtain lift and then applied them to artificial gliders with the connection of reinforcement learning.^{30,31} In 2018, Verma *et al.* used a DRL framework to explore how fish can obtain the most energy-saving formation based on the vortex generated during the fish colony parade.³² Compared with traditional research methods, the RL learns a more effective and direct mapping relationship, and reveals the internal mechanism of the strategy, consistent with the demand and nature of the AFC. In 2019, Belus and Rabault maintained the stability of the one-dimensional depth-integrated falling liquid film based on the AFC framework with DRL.³³ Then Rabault modeled and studied the two-dimensional flow around a cylinder with Reynolds number 100 and used the well-known proximal policy optimization DRL algorithms to control the symmetrical nozzles on the cylinder poles with the aim of resistance reduction.³⁴ It was found that CFD took up most of the training time, so a multi-environment parallel model was built to accelerate the training process.³⁵ Considering that DRL owns a good generalization ability, Rabault extended their work by conducting

research in detail with the Reynolds number between 60 and 400.³⁶ In addition to the powerful ability of the DRL algorithm, the CFD accuracy of the environmental model and timely transmission of the sensors are both the important reasons for the success of AFC. Paris *et al.* combined sparsity-seeking algorithm with DRL algorithm to optimize the layout of sensors, which highlighted possibilities of DRL and paved the way to efficient, robust, and practical implementations of these control techniques in experimental or industrial system.³⁷

The main objective of current work is to search for a robust active flow control strategy based on a pair of symmetrical jets to suppress the vortex-induced vibration of a cylinder in the flow at low Reynolds number of $Re = 100$. There are two effective machine learning methods used to find the optimal active flow control strategy, including AL and DRL. In order to discuss the association, two basic models are introduced, one of which, named the agent model, is responsible for decision-making and internal iterative learning. The other model environment is responsible for the numerical simulation of the physical environment and feedback for the control action. In the following, the physical problem and model setup methodology are described in Sec. II. The results of AL agents and DRL agent are presented, analyzed, and discussed in Sec. III. Finally, the conclusions and expectations are drawn in Sec. IV.

II. ENVIRONMENT MODEL AND METHODOLOGY

A. Environment model

In the present work, a computational fluid dynamics method is adopted to establish the fluid environment, and the physical model is a two-dimensional case of vortex-induced vibration of a cylinder with Reynolds number of 100. As shown in Fig. 1, there is a fixed rigid cylinder immersed in the incompressible flow, the diameter D of the cylinder is 1 m, which is set as the reference length. The downstream length of the whole flow field is 29D, and the transverse width is 16D.

The distances from the cylinder center are 8D to the upper and lower wall, 8D from the front inlet and 21D from the rear outlet. The flow configuration is set as follows: the inlet velocity is constant 1 m/s, the fluid density ρ is 1 kg/m³, and the kinematic viscosity coefficient μ is 0.01 kg/(m s).

The numerical simulation of the flow is performed on the open source software OpenFOAM.³⁸ The PIMPLE algorithm³⁹ is adopted for the couple of the pressure and velocity equations. We use the second-order implicit backward scheme for the time discretization. For the spatial discretization, the Gauss linear scheme is adopted for gradient term discretization, the cell-limited second-order Gauss linear upwind scheme is adopted for velocity diffusion term discretization, and the corrected Gauss linear scheme is adopted for Laplace term discretization, respectively. This study relies on the six-degree-of-freedom solver (though single degree of freedom is considered in this paper) and the dynamic mesh technology to simulate the vortex-induced vibration case. Considering the movement of cylinder boundary, the arbitrary Lagrangian–Eulerian method⁴⁰ discretizes the conservation equation for arbitrary control volume V , which is expressed as

$$\frac{\partial}{\partial t} \int_V \rho \mathbf{U} dV + \oint_S \rho \mathbf{U} (\mathbf{U} - \mathbf{U}_s) \cdot d\mathbf{S} = \oint_S \rho \nabla(\nu \mathbf{U}) \cdot d\mathbf{S} - \int_V \nabla p dV, \quad (1)$$

$$\frac{\partial}{\partial t} \int_V dV - \oint_S \mathbf{U}_s \cdot d\mathbf{S} = 0, \quad (2)$$

where S is the boundary surface of the control volume, \mathbf{U}_s is the velocity of the moving surface, and \mathbf{U} is the velocity of flow. ρ is the density, p is the pressure, and ν is the kinematic viscosity.

The fact that the vortices alternately shed from both sides of the elastic cylinder surface produces a periodic pulsating force in transverse direction, leading to the periodical vibration of the cylinder. The Reynolds number is an important dimensionless number, which is

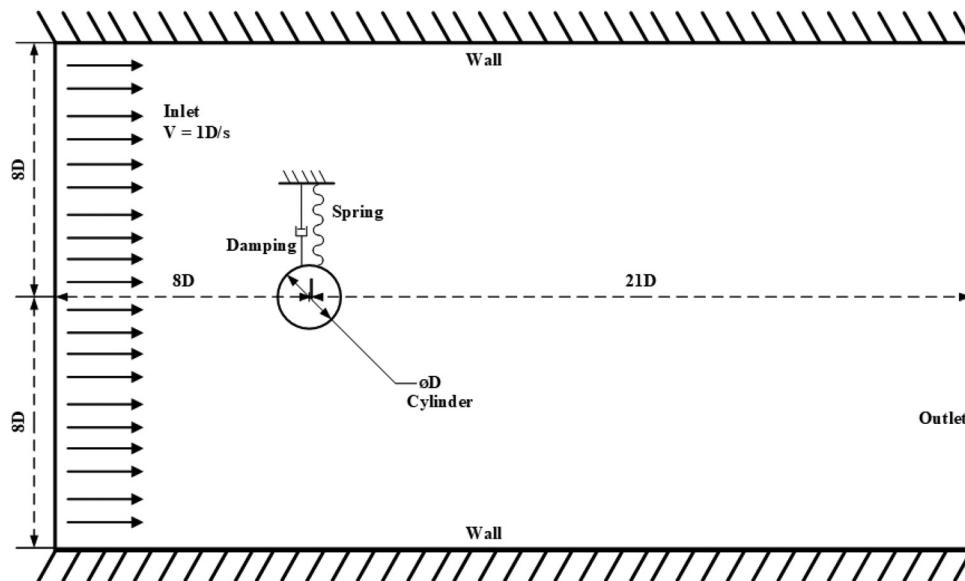


FIG. 1. Computational domain and definition of flow over the cylinder at low Reynolds number of $Re = 100$. The cylinder is vibrating induced by the vortex only along the crossflow direction, whose motion is constrained by damping and spring.

used to characterize the viscous effect of flow, especially the vortex mode on the wake.⁴¹ The Reynolds number is set to 100 in our case, which is expressed as

$$\text{Re} = \frac{\rho UD}{\mu}, \quad (3)$$

where U is the inflow velocity, D is the reference length, ρ is the density of the flow and μ is the viscosity coefficient of the flow.

In order to simplify the problem, we only focus on the motion of the cylinder in the transverse direction, and the second-order motion equation of the cylinder is governed by

$$m\ddot{y}(t + \Delta t) + \xi\dot{y}(t + \Delta t) + ky(t + \Delta t) = F(t + \Delta t), \quad (4)$$

where \ddot{y} is the acceleration in the transverse direction of the cylinder's centroid, \dot{y} is the velocity, F is the fluid force on cylinder, m is the mass of the cylinder, ξ is the damping coefficient of the damper, and K is the elastic coefficient of the spring. To solve this motion equation, Newmark method based on linear acceleration assumption is adopted in this study.

Different masses m , damping coefficients ξ , and elastic coefficients K correspond to different branches and wake vortex structures. Usually, a combined dimensionless parameter, the reduced velocity (U_r), is used to reflect the correlation with the vibration form of cylinder, which is expressed as

$$U_r = \frac{U}{f_{nat}D}, \quad (5)$$

where U is the inflow velocity, f_{nat} is the natural frequency of the mass-damping system, and D is the reference length.

We have checked the grid scale and time step before we start to explore the active flow control strategy by numerical simulations. Four sets of grids are checked on the vibration amplitude and Strouhal number. The Strouhal number is a dimensionless parameter characterizing unsteadiness effect of flow, which is expressed as

$$Sr = \frac{Df}{U}, \quad (6)$$

where U is the inflow velocity, f is the periodic motion frequency of the VIV cylinder, and D is the reference length.

With the increase in the mesh quantity, the stable vibration amplitude and the Strouhal number of flow gradually converge, as listed in Table I. Except grid M1, the vibration amplitudes of the other grids are very close, which are acceptable. Due to the need to collect a large number of data, the grid M2 is chosen. The time step has the

TABLE I. The effect of grid size on simulation results.(bare cylinder, $Ur = 4.85$, time step = 0.005, 8 cores, Intel(R) Xeon(R) Platinum 8175M Central Processing Unit (CPU) @ 2.50 GHz).

Size	Total cells	Total points	Amplitude	Sr	Consuming time (s)
M1	6 126	6 251	0.1857	0.1905	1361.79
M2	11 398	11 598	0.5948	0.2018	2942.43
M3	14 564	14 800	0.5944	0.2016	3791.91
M4	17 619	17 884	0.5941	0.2016	4753.66

TABLE II. The effect of time step on simulation results.(bare cylinder, $Ur = 4.85$, grid M2, 8 cores, Intel(R) Xeon(R) Platinum 8175M CPU @ 2.50 GHz).

Time step	Amplitude	Sr	Consuming time (s)
0.1	3.76	0.1923	191.89
0.05	1.295	0.1923	346.38
0.01	0.6466	0.2174	1 609.62
0.005	0.5948	0.2018	2 942.43
0.001	0.5645	0.2015	10 674.03
0.0005	0.5621	0.2174	18 174.28

greatest influence on the consuming time of numerical simulation with fixed simulation time. Six time steps are checked in the same numerical simulation with fixed simulation time equal to 200 s as listed in Table II. It is obvious that the convergence amplitude of cylinder is around 0.56, in which case only time step 0.001 and time step 0.0005 are eligible, but the consuming time of both is extravagant. The amplitude of the case with time step 0.005 has a 6.7% error compared to 0.56, which is nearly acceptable. Moreover, we have to pay nearly three times the computing time for the 6.7% error if time step 0.001 is used. Based on the above considerations, the grid M2 is chosen and the time step is set as 0.005.

For further validation, the schematic diagram of reduced velocity and amplitude as shown in Fig. 2 is drawn and compared with the results with other literatures. The dimensionless parameter A_Y^* associated with the cross-flow amplitude appears frequently, which is expressed as follows:

$$A_Y^* = \frac{A}{D}, \quad (7)$$

where D is the diameter of the cylinder and A is the amplitude of cylinder in the cross-flow direction. In this study, the reference length, i.e., the diameter of the cylinder, is set to 1 m. Therefore, instead of A_Y^* , amplitude A is used directly in the following part of the article. In Fig. 2, it is seen that the blue curve is basically consistent with our curve in the initial branch, which has the same flow configuration.

The objective of our research is to suppress vibration and to reduce drag of the cylinder. The vibration of the cylinder is solved by Eq. (4), and the force of the cylinder is obtained through numerical

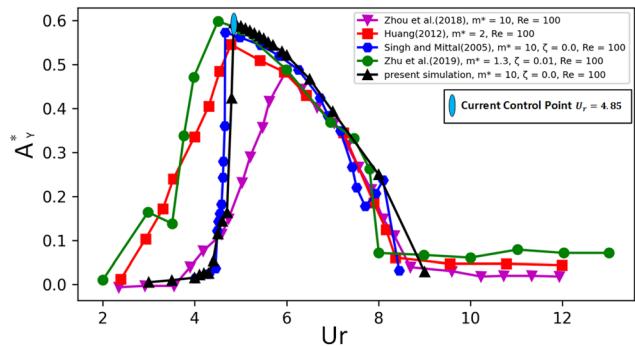


FIG. 2. Model validation: oscillation amplitude of the bare cylinder vs the reduced velocity.²⁰

simulation. Two non-dimensional parameters, C_L and C_D , are defined for the quantification of objectives as follows:

$$C_L = \frac{\int_C (\sigma \cdot n) \cdot e_y \, dS}{\frac{1}{2} \rho \bar{U}^2 D}, \quad (8)$$

$$C_D = \frac{\int_C (\sigma \cdot n) \cdot e_x \, dS}{\frac{1}{2} \rho \bar{U}^2 D}, \quad (9)$$

where σ is the Cauchy stress tensor, n is the unit vector normal to the outer cylinder surface, C is the surface of the cylinder, D is the diameter of the cylinder, ρ is the volumetric mass density of the fluid, \bar{U} is the velocity of the uniform flow, $e_x = (1, 0)$, and $e_y = (0, 1)$.

The active flow control work in this study will be carried out under the condition where U_r is 4.85 when the vortex shedding frequency is near the natural frequency. Our AFC framework is divided into two parts including environment part and agent part. Two different frameworks are transplanted into the establishment of agents including AL and DRL, which will be introduced in Secs. II B and II C. Environment needs input of control action offered by agent and outputs simulation results by CFD platforms. There are two actuator jets placed on two poles of the cylinder, which have the parabolic spray-velocity spatial profile with a smoother transition to the cylindrical surface. The sum of the control actions of upper and lower nozzles is set to zero to avoid extra storage space. The flow information collected by sensors around the cylinder will be transferred to the agent as observation if needed. Dynamic parameters (velocity, force, displacement, etc.) are the objectives of the AFC, which are calculated on the OpenFOAM. Through the interaction between agent and environment, the agent adjusts control strategy approaching the goal. For the sake of the limitation of jet power in practical application and the

exploration cost on large action space, the given action value is not allowed to be larger than 3.0. The AFC framework is shown in Fig. 3.

B. Active learning agent framework

This section introduces an active learning agent framework, which we use to establish the first agent model. As a branch of machine learning, active learning is mainly designed for the scenario where data tags are few or the cost of labeling is high. The AL agent framework sets up a GPR surrogate model to explain the relationship between control action and the vibration state of the controlled cylinder. The role of Bayesian optimization is to efficiently sample the best point to achieve a locally accurate model surrogate model considering the balance between exploration and exploitation. In this work, the input of the surrogate model is the parameter set of a specified control function that is selected from the function library. With the demand of practical application, the output of GPR surrogate model is the stable vibration amplitude of VIV under control.

As a representative of the nonparametric method, the GPR may not have the computation speed of the parametric method, but it does well in training efficiency. With assume that all sampling points satisfy the Gaussian distribution, the Gaussian process is a random process, which can be uniquely defined by a mean function $\mu(x)$ and a covariance function $\kappa(x_i, x_j)$. Given a training dataset at the input sampling points X and the corresponding label values f , a Gaussian process prior can be converted into a Gaussian process posterior according to Bayesian theory, which is used to predict the prediction label f^* at the test point X^* ,

$$f^*(X^*) \sim N(\kappa(X^*, X)\kappa(X, X)^{-1}(f - \mu(X)) + \mu(X^*), \kappa(X, X) - \kappa(X^*, X)\kappa(X, X)^{-1}\kappa(X, X^*)). \quad (10)$$

The covariance function κ , also known as kernel function, is used to characterize the correlation between different random variables, such

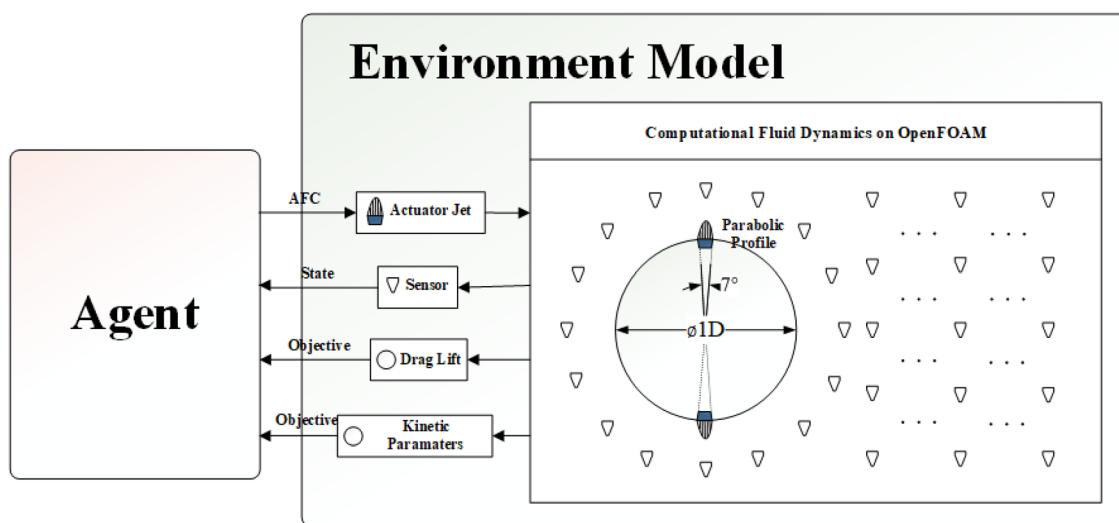


FIG. 3. Schematic of the environment model and the active flow control framework. A pair of jets symmetrically is mounted on the pole of the cylinder, and 152 sensors are placed on the wake of the cylinder to observe the flow state. After one numerical simulation on the new control action, the environment provides the latest flow states and the corresponding dynamic information (velocity, force, displacement, and so on) to the agent.

as, the smoothness of continuous measurement data. The Matern kernel function is used in this work, whose superparameters can be determined by the maximum likelihood criterion.⁴²

The GPR surrogate model establishes a map from open-loop control parameters to control objectives. Due to the expensive of numerical simulation, there is no need to fully explore the action space for active learning agent. Hence, Bayesian optimization method with its specific sampling rule is introduced for its balance between exploitation and exploration, which means that only most valuable points (the optimal or the most uncertain solutions in the current iteration) are sampled. This sampling rule in Bayesian optimization is the acquisition function EI (expected improvement), whose extreme point is the next sampling point,

$$\max_x EI(x|X), \quad (11)$$

$$EI(x|X) = \begin{cases} (\mu(x) - f(x^+) - \xi)\Phi(z) + \sigma(x)\phi(x) & \text{if } \sigma(x) \geq 0, \\ 0 & \text{if } \sigma(x) < 0, \end{cases} \quad (12)$$

$$z = \begin{cases} \frac{\mu(x) - f(x^+) - \xi}{\sigma(x)} & \text{if } \sigma(x) \geq 0, \\ 0 & \text{if } \sigma(x) < 0, \end{cases} \quad (13)$$

where x^+ and $f(x^+)$ represent the best objective function point and value given by the surrogate model (since the surrogate model has no consumption in calculation, the Broyden-Fletcher-Goldfarb-Shanno method (BFGS) method⁴³ is used for optimization search in this study), $\mu(x)$ and $\delta(x)$ represent the mean and variance at point x , Φ

and ϕ are the probability density function and cumulative distribution function of standard normal distribution. Here, a new parameter ξ is adopted to control the trade-off between exploitation and exploration, whose larger value means more exploration and smaller value means more exploitation. At first, 30 points by Latin hypercube sampling is prepared for the initial surrogate model. The episode is defined as one simulation from the start point to the time when the cylinder with one specified control function gets to its ultimate periodic oscillation. For each episode, the Bayesian optimization does not send one most valuable point to the environment, which returns a stable vibration amplitude by numerical simulation, until reaching the maximum number of episodes. The decision-making agent model based on Gaussian process regression and Bayesian optimization framework is shown in Fig. 4.

C. Deep reinforcement learning agent framework

In this section, we present the establishment of the deep reinforcement learning framework with the core of the soft actor-critical algorithm (SAC).⁴⁴

Reinforcement learning is one of the linchpins in machine learning, which comes from behaviorism theory in psychology. On the basis of external reward or punishment, it gradually forms the expectation of stimulation and produces the habitual behavior to obtain the maximum benefit.

Different from the active learning framework, the DRL establishes an interactive decision-making agent, which receives the information from the sensor placed around the cylinder, fully observes the state of the environment, and then outputs a control action to the actuators. Therefore, what the agent gives is the optimal action a_t with

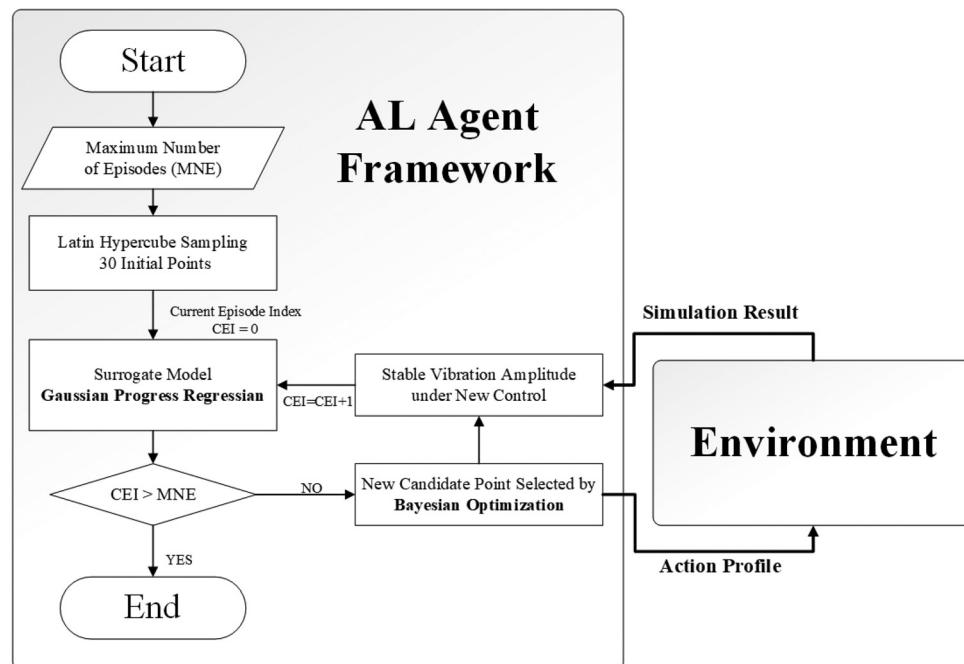


FIG. 4. Schematic of agent establishment process in the active learning framework. A surrogate model based on the Gaussian process regression method is established as a view of control action space. The Bayesian optimization method is used to improve model accuracy in the least samples from expensive environment simulations. As open-loop method, the optimal parameter set of the AL agent is not decided until maximum number of episodes is reached.

the greatest instantaneous reward R_t under current observation state S_t . The optimal solution is not related to earlier observation but only current observation S_t , which is also called the Markov characteristic. The significant instantaneous reward R_b , which determines the direction of strategy optimization, consists of a target revenue and an action penalty. The target revenue is about the motion (position, velocity) and force (drag, lift) coefficients of the cylinder, and the action penalty is comprised of the bias (action absolute value) and variance (difference between adjacent actions) of the control action,

$$a_t = \alpha force_t + \beta motion_t + \gamma bias_t + \delta variance_t, \quad (14)$$

where α , β , γ , and δ are the ratios of four segments, and t is the time index.

The task of establishing control strategy π is completed by three-layer artificial neural networks (ANNs), known as actor. The input of artificial neural network is the observation of flow, and the output is an action distribution. The reason why the control chooses a stochastic strategy is to avoid trapping into local optimal solution. Except the actor, SAC algorithm framework sets up a critic composed of ANNs as well. The role the critic plays is to evaluate the score which the action a_t will get with the observation S_t . The higher value means a greater probability to take such an action. In actor-critic framework, the critic acts as such a significant supervisor that four artificial neural networks are used in the meantime to establish a critic; each has two hidden layers with 128 neurons. Two of them are target networks and the left are current networks. With the idea of regression, the current evaluate value $Q(S_t, a_t)$ from current networks is updated with instantaneous reward and future expectation from target networks. Smaller output of two same-type neural networks is taken for the consideration of radicality and convergence.⁴⁵ The activation function of all hidden neurons is the ReLU function.

The ultimate objective of typical DRL is to maximize the cumulative of instantaneous revenue from the initial time to the end time, but SAC algorithm used in current study has the idea of maximum entropy H of policy, so as to randomize the probability of output dispersed as much as possible rather than concentrated on one current-optimal action. The introduction of information entropy gives a greater exploration mandate to the agent, resulting in a strategy with stronger robustness. Especially when the actual environment value function presents multiple peaks, the strategy of SAC algorithm can get rid of local single optimal solution, which has a great advantage in solving the multi-mode problem. The optimization objective of SAC algorithm is given by

$$\pi^* = \arg \max_{\pi} E_{(S_t, a_t) \sim \rho_{\pi}} \left[\sum_t R_t(S_t, a_t) + \alpha H(\pi(\cdot | S_t)) \right], \quad (15)$$

where ρ_{π} is environment state transition probability and α is the ratio of entropy. SAC algorithm is an off-policy episode-based DRL algorithm, which means the agent gains experience and stores the data into a buffer in every episode. Similarly, the maximum number of scenes specified is used as the condition for the end of training. One 50 s episode is adopted, which is different from the flexible-time episode of active learning framework. The agent interacts with the environment 200 times per episode and trains itself with the sampled data from buffer. The whole decision-making agent model based on soft actor-critic deep reinforcement learning framework is shown in Fig. 5.

III. RESULTS AND DISCUSSION

A. Active flow control with active learning agent

In active learning framework, we choose a specified function from the pre-designed function library to establish a GPR surrogate model. The Bayesian optimization method is then used to improve the accuracy with the least sampling times and find the optimal control parameters in the action space. Without loss of generality, the beginning of each episode is set to the point when the cylinder is oscillating upward at the equilibrium point.

In view of the unknown high-dimensional and highly nonlinear fluid structure coupling dynamics problems, there is no prior knowledge to decide which type of mathematical function has an ability to eliminate the vibration. Three basis functions are selected, including constant function [invariance, Eq. (16)], exponential function [monotonicity, Eqs. (17) and (18)], and triangle function [periodicity, Eq. (19)] as

$$Action_t = C, \quad (16)$$

$$Action_t = A \cdot e^{\frac{\alpha}{coe} t}, \alpha = \frac{\log(0.618)}{50}, \quad (17)$$

$$Action_t = A \cdot (1 - e^{\frac{\alpha}{coe} t}), \alpha = \frac{\log(0.618)}{50}, \quad (18)$$

$$Action_t = A \cdot \cos\left(\frac{2\pi}{T}t + \phi\right), T = 4.955, \quad (19)$$

where $action_t$ is the current action, A is the amplitude, coe is the coefficient, ϕ is the phase, and t is the current time.

Constant function is one of the most common control functions, with a constant injection or suction speed of both set to the jets from begin to end. The iterative learning process of AL agent is shown in Fig. 6. Ten-point-set is sampled for the use of initial construction. It is almost enough that the ten initial points generated by one-dimensional hypercube Latin sampling to capture the influence of different constant values on vibration amplitude. The results show that the constant control function can suppress the vortex-induced vibration when the upper jet is in suction mode with the lower injection jet or the upper jet is in injection mode with the lower suction jet. The suppression strengthens with the increase in the absolute value, until the constant action reaches 2.6. Since the example of constant is not complicated, BO is stopped after only five iterations. The lowest controlled amplitude is 0.5412 with a positive equilibrium offset of 0.066.

The use of exponential function takes the influence of monotonicity of control action on vibration reduction into account. Two exponential function forms are adopted, which are uniquely represented by the parameter amplitude and coefficient. For the calculation operability, a numerical transformation is applied to the parameter coefficient to cover most meaningful control strategies, which follows the golden ratio. The two-dimensional GPR surrogate model based on 20 initial points forms a horseshoe shape distribution in the explored action space. No matter which kind of monotonicity, the larger control amplitude always has a better suppression effect as shown in Fig. 7. Moreover, according to the choice of coefficient, the function with faster growth will be preferred. Similar to constant control, the exponential control agent has a preference for large action amplitude. The amplitude is controlled at about 0.545 in the form of both exponential

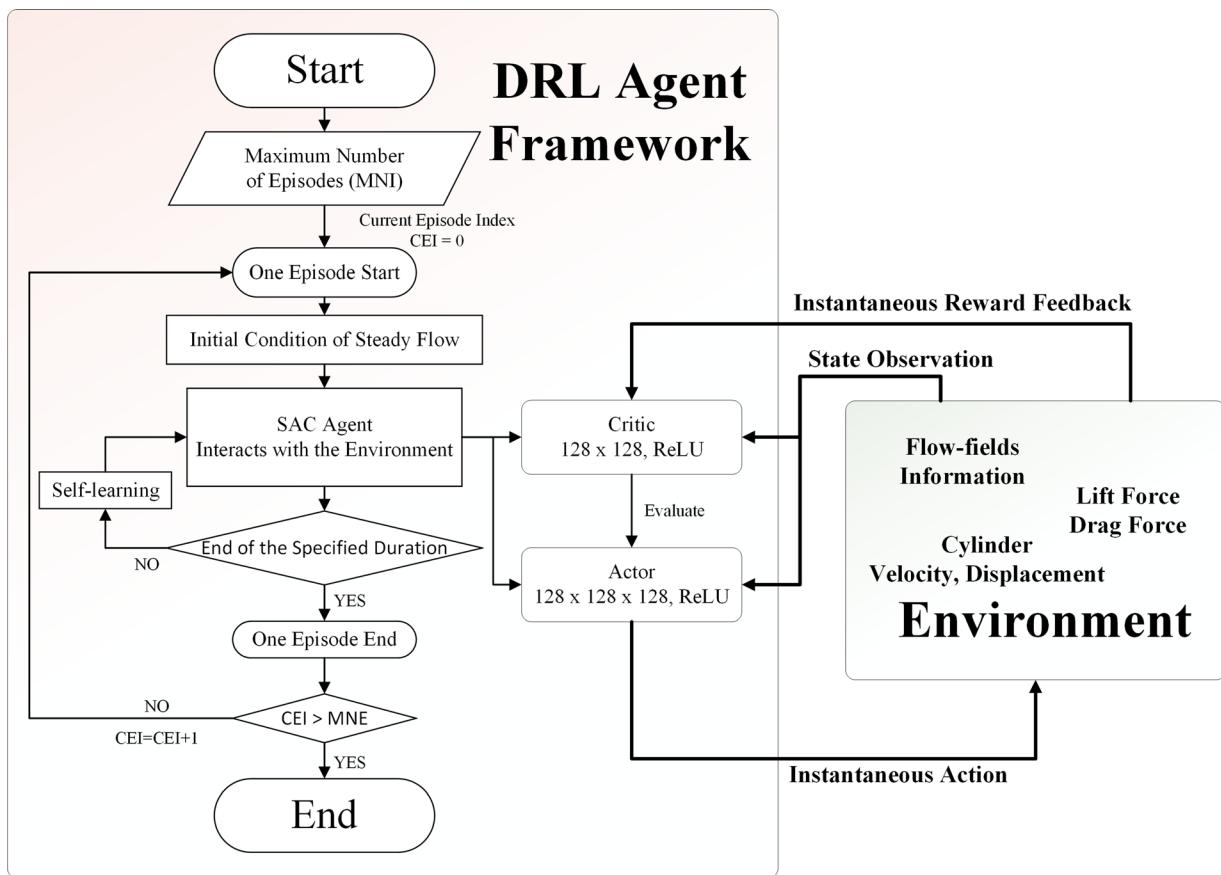


FIG. 5. Schematic of agent establishment process in the soft actor-critic deep reinforcement learning framework. The role of critic is to evaluate the value of the control strategy taking pre-designed instantaneous feedback, which is offered to actor to output the action with more revenue on current observation flow state. The control on DRL is closed-loop, which needs a lot of interactions between the agent and environment every step to train the artificial neural network iteratively. Similar to the GPR-BO framework, the maximum number of episodes is given as a sign of end.

function and constant function, which is only 9.17% lower than the amplitude of 0.60 without control.

In VIV, such a periodic problem, periodic control strategy with a pair of transverse jets, is also explored. In Eq. (19), the period T is directly set as the natural frequency (near the frequency of the cylinder oscillation) of the cylinder so as to avoid unnecessary vibration noise. Similar to exponential, 20 points consisting of the amplitude and phase ϕ are sampled with two random parameters in advance for the initial use. The output is the amplitude of stable vibration, as shown in Fig. 8. The three-dimensional diagram exhibits a complex model with valleys and peaks. The optimal solution is not exposed directly like previous cases. It is necessary to search thoroughly the low amplitude region represented by blue with the aid of the Bayesian optimization. At the same time, the area where the model uncertainty is relatively large should be included in avoidance of the local optimal solution, which is determined by the EI function. After a total of fifteen optimization iterations, the model identifies the best parameter set ($A = 0.8$, $\phi = 1.35$) and the ultimate amplitude is 0.43 with a descent of 28.3%.

Based on the past experience of constant and exponential, the action near its boundary (3.0) is the place good control strategies

gather around. However, the GPR model tells us that the best control amplitude is quite a small action, which makes it a confusing point. We do a comparison between the large-action (defined as $A = 3.0$, $\phi = 1.35$) and small-action (defined as $A = 0.8$, $\phi = 1.35$) control, and the result is shown in Fig. 9. It is shown that the development of amplitude is not monotonic with a minimum amplitude = 0.189, which is the lowest temporary amplitude. The stable amplitude controlled by the large-action triangle agent is 0.746, which is even higher than the stable uncontrolled amplitude. The reason why the large-action agent shows a strange motion mode and the way to meet it with low-amplitude vibration will be discussed in Sec. III C.

B. Active flow control with deep reinforcement learning agent

The small-action active flow control with the AL agent has already been able to reduce the amplitude of vibration in the form of regular function. As a powerful intelligent control tool, deep reinforcement learning has an ability to find a relatively robust strategy especially in highly dimensional and nonlinear systems, which has arisen

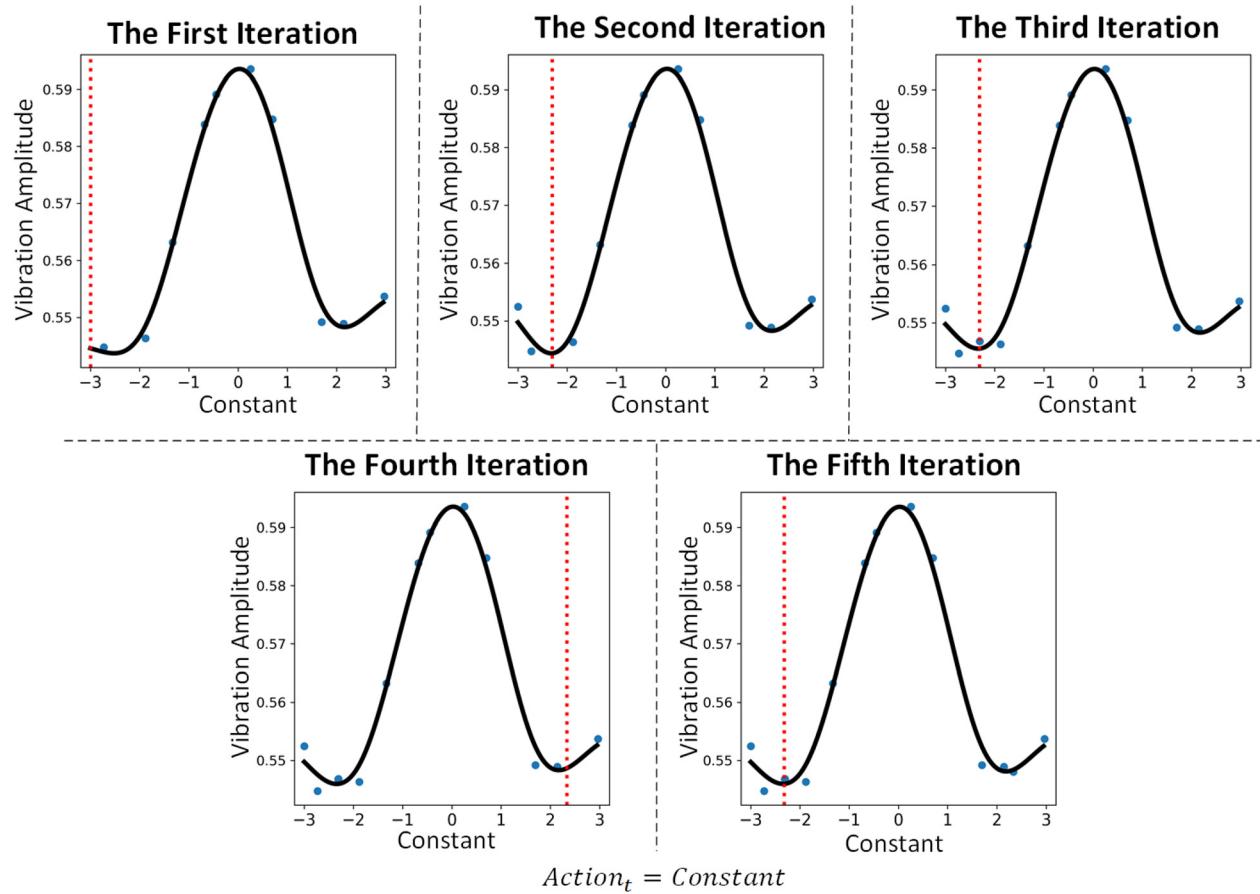


FIG. 6. Iterative process of constant control function. 10 sample points are used to establish an initial GPR agent. The black solid line is the prediction of current agent on actions between -3 and 3 , and the red dotted line is proposed by the BO method, which means a new urgent sample point. The blue points are the former samples, which are expensive simulation results. The action of each moment is given by the formula at the bottom of the picture.

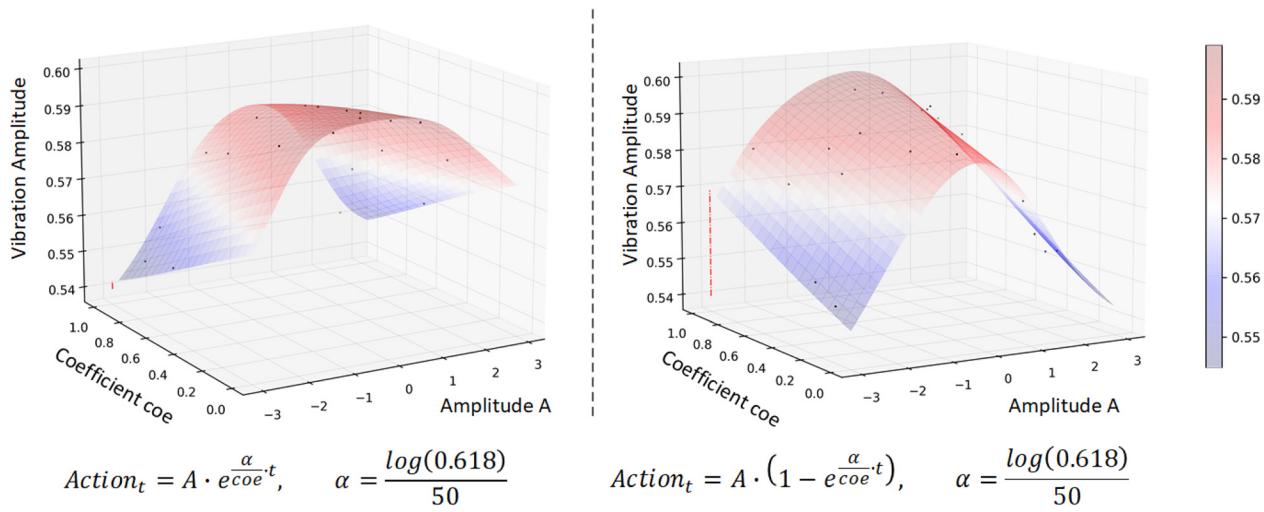
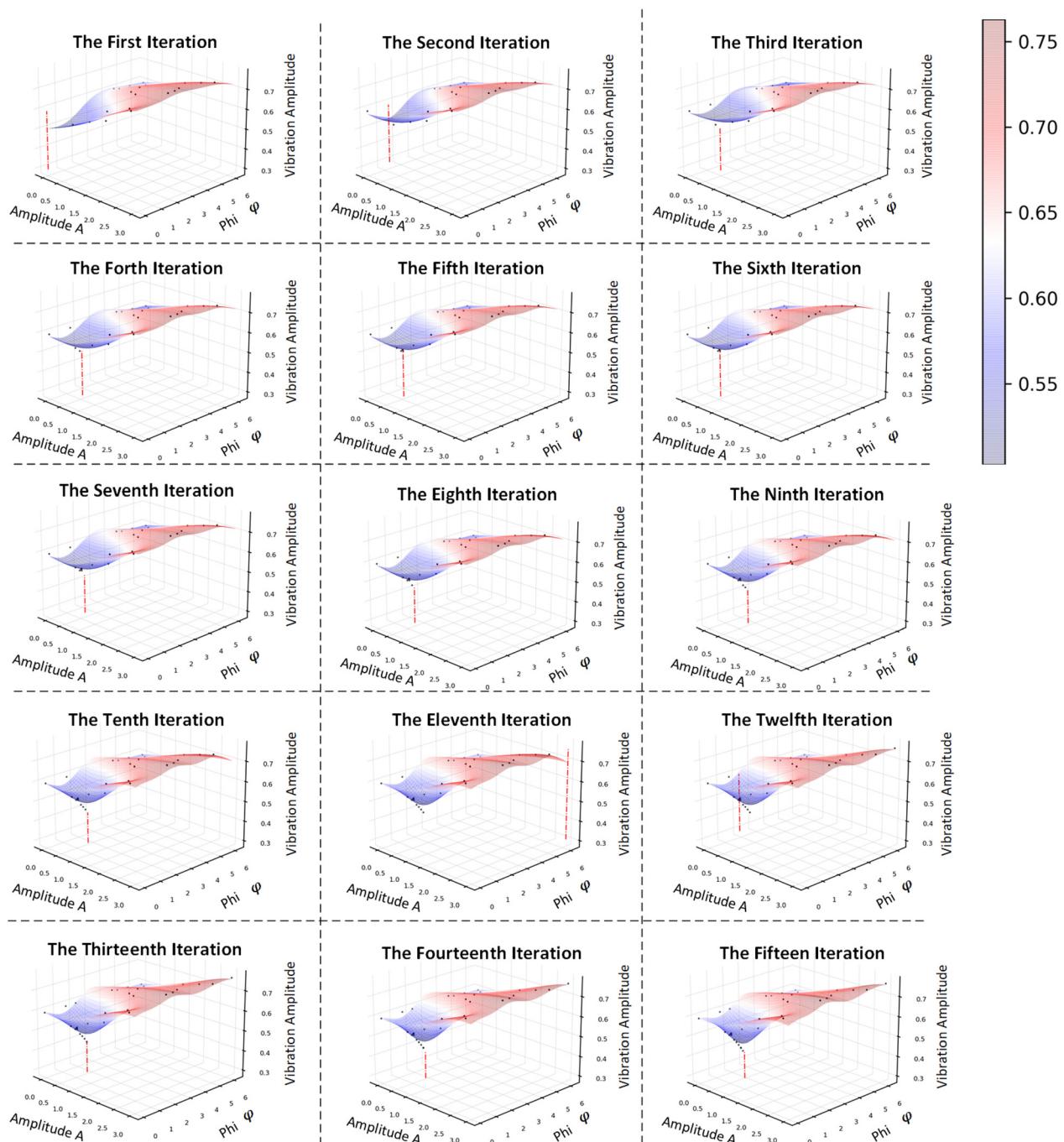


FIG. 7. Ultimate predicted results of exponential control functions in 20 sample points. The dotted dashed line means next sample points, which is unnecessary to continue optimization for an obvious optimal solution at amplitude = -3 . The black points are the former samples. The action of each moment is given by the formula at the bottom of the picture.



$$Action_t = A \cdot \cos\left(\frac{2\pi}{T}t + \varphi\right), \quad T = 4.955$$

FIG. 8. Iterative process of triangle control function. Initial GPR triangle agent sample 20 points and another 15 optimization iterations are used to search around. Each red dot dash line means the most valuable sample point provided by EI function in optimization. The blue valley of the surface has the most intensive exploration, which means the location of the minimum amplitude of vortex-induced vibration. The action of each moment is given by the formula at the bottom of the picture.

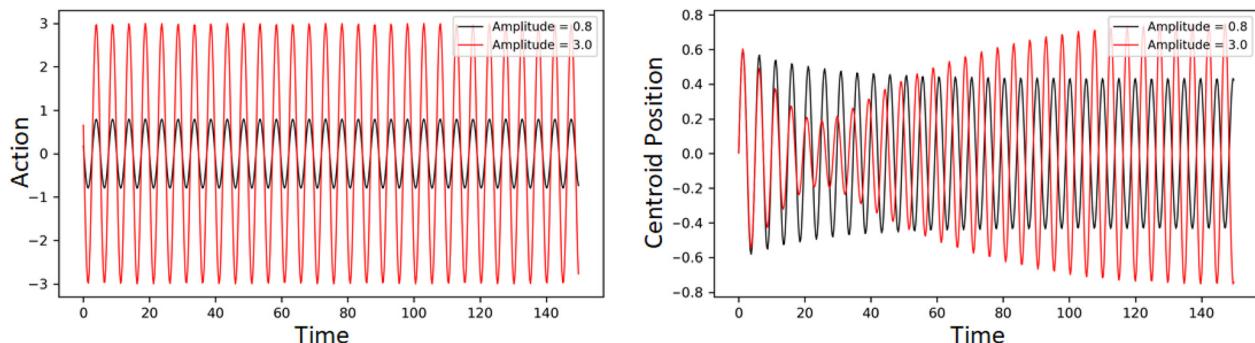
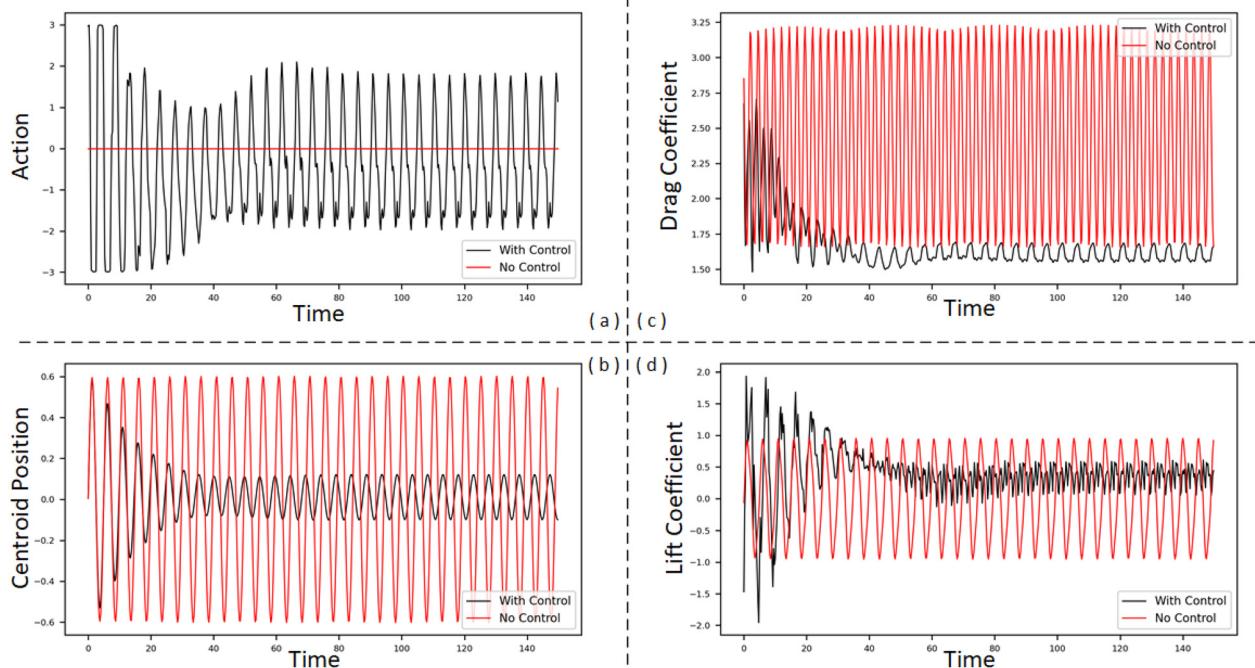


FIG. 9. Trajectories of large-action triangle AL agent and small-action AL agent. (a) The trajectory of action shows different amplitude but same frequency in triangle function form. (b) The position trajectory controlled by large-action AL agent shows a trend of falling first and then rising. Small-action triangle AL agent can slowly suppress the vibration to a certain extent.

more expectations. Known as a closed-loop control system, the agent gives an instantaneous action by receiving the information collected by the sensors placed in the flow. After processing the control signal, the environment gives a corresponding feedback, which is used to modify the parameters of the deep neural network in the agent, rather than the parameters of the control function. Benefit from the extraordinary ability of fitting nonlinear function of deep neural network, it is possible for DRL to gain a periodic nonlinear control strategy with a

stronger suppression effect. For comparison, the beginning of each episode is the same as the active learning framework.

After about 800 episodes of training, the weights of the neural networks are extracted and applied to a new test agent. The greedy action is adopted in every decision to test the optimal solution in the current situation. Different from the open-loop control function, the SAC agent outputs a concrete action according to the feedback from the environment, as shown in Fig. 10. It is obvious that this unique



$$\text{Action}_t = \text{SAC Agent}(\text{Actor}, \text{Critic}, \text{State}_t)$$

FIG. 10. Trajectories of the DRL-controlled test case by black and the uncontrolled case by red. (a) The DRL agent gives a strong time-varying periodic trajectory of the action, which keeps oscillation with an amplitude = 2 in the end. (b) The position vibration amplitude of the centroid of the control cylinder decreases rapidly with time and finally stabilizes around 0.11, which decreased by 82.7%. (c) The controlled time trajectory of drag coefficient shows a sharp decline, and the stable time-mean drag coefficient is 2.035, which is 17% lower than the uncontrolled case. (d) Trajectory of lift coefficient. The controlled lift shows an offset to zero due to the influence of asymmetric action.

periodic control law is almost impossible to be expressed by any known expression. The control process of agent can be divided into four stages. The first stage is from 0 to 15 s, during which the transition between actions is very smooth, marked by large-amplitude action strategy. The vibration amplitude of the cylinder decreases fastest. The amplitude of action decreases from 3 to 1.4 continuously in the second stage from 15 to 45, and the vibration equilibrium point is no longer at zero. The vibration amplitude is still decreasing, but the decreasing speed slows down to the lowest point of 0.08. In the third stage from 45 to 75 s, the movement amplitude rises, especially with a sharp change of the maximum value of the action. There is a slight increase on the motion amplitude, which is not much obvious. When the amplitude of the action vibration remains about 2.0, and the average value is near the origin, it comes to the last stage. The action is more radical even with some fluctuations at the pole. The stable vibration amplitude maintains at 0.11 and the mean position is around 0.01, which is quite near the original equilibrium point. Compared with the uncontrolled vibration case, the active flow control led by SAC agent successfully suppressed the vibration amplitude below 0.11, which is decreased by 82.7%.

The most noticeable variable is vorticity in this case, which gives rise to the generation of VIV. Therefore, the key to eliminate vibration is to restrain the increase in vorticity. As shown in Fig. 11, we capture the instantaneous vorticity field under three control conditions. When the uncontrolled cylinder reaches the highest point, the vortex II from the lower surface becomes strongest. Most of the surface of the cylinder is occupied by the positive vortex, which is characterized by red, while the blue vortex I is in the state of separation. When the cylinder moves to the lowest point, the red vortex II sheds off and the new vortex III on the upper surface grows strong, and the vortex I completely falls off.

When the cylinder moves to the top just with a start of control, the jet on the lower surface of the cylinder is injecting in the maximum power, mixed with the flow through the surface to form a counter rotating fluid so as to decline the power of vortex II. On the opposite surface, the work of the jet is to suck in the tail fluid of vortex I, which is leaving. Reaching the lowest, it is found that the new vortex III is disturbed by the injection effect of the jet on the upper surface as well, with suction on the lower surface.

If the cylinder under active flow control vibrates with a small amplitude, the interference mode of the agent is different. Whether at the highest position or the lowest, the action amplitude is relatively small at this time. The uncontrolled vortex I is pulled very long in the transverse direction due to the strong shear effect and the stable controlled vortex I is very compact directly behind the cylinder on the contrary. For the vortex II at the highest or the vortex III at the lowest point, they have lost the power to occupy the majority of the wall, which is also caused by the suppression of vorticity.

Figure 12 shows the average vorticity of the flow in one stable vibration period, and it is seen that the distribution of vorticity has changed greatly under the active flow control. In the typical time-averaged vorticity distribution diagram of vortex-induced vibration, there are four visible vortices symmetrically placed in the wake of the cylinder. The first two vortices are behind the cylinder, and the last two vortices with smaller energy are farther downstream. Under control, the vortices hidden in the vortex street that were originally shed off are also closely united near the centerline of the flow field. The maximum width of vortex street is 4.08 m, while that of the controlled vortex street is only 2.75 m, with a decrease in 32.6%. More obviously, the color of the controlled vortex trace by blue in the downstream region has faded, which strongly proves the attenuation of the shedding vortex strength.

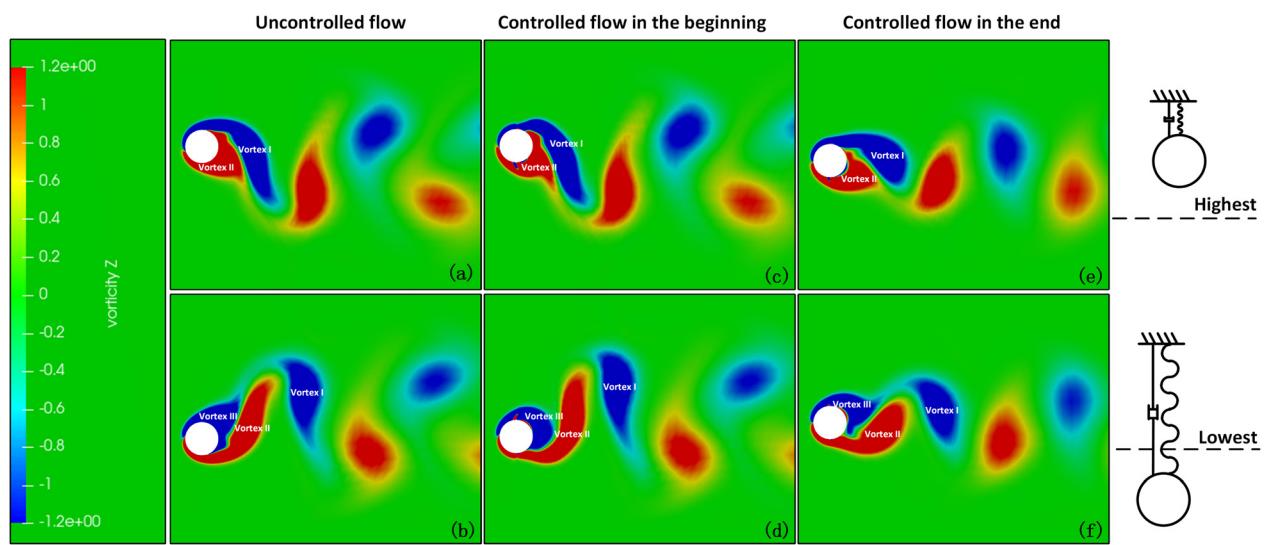


FIG. 11. Color-coded vorticity contours. (a) Snapshot of the vorticity contour when the uncontrolled cylinder reaches its highest positon in crossflow direction. (b) Snapshot of the vorticity contour when the uncontrolled cylinder reaches its lowest positon in crossflow direction. (c) Snapshot of the vorticity contour when the controlled cylinder reaches its highest position in the beginning. (d) Snapshot of the vorticity contour when the controlled cylinder reaches its lowest position in the beginning. (e) Snapshot of the vorticity contour when the controlled cylinder reaches its highest position in the end. (f) Snapshot of the vorticity contour when the controlled cylinder reaches its lowest position in the end. The vortex that occupies the majority of the surface like vortex II at the highest position and vortex III at the lowest position is disturbed by ejection, and the vortex that is shedding off is inhaled. In the motion of stable vibration, the vortex has lower power and stays closer to the cylinder in crossflow direction.

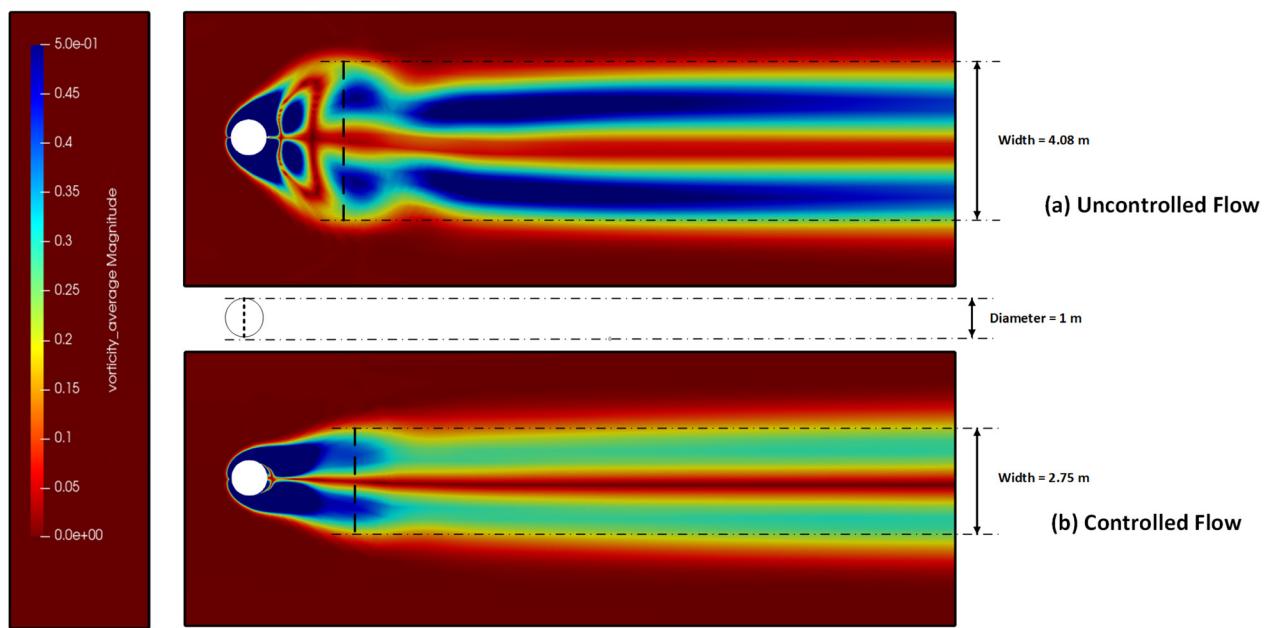


FIG. 12. Vorticity periodic time-average contour. The controlled vorticity contour in (b) shows a different shedding phenomenon compared to uncontrolled case in (a), which has a narrower vorticity street and less rotation strength.

The suppression of vibration needs to change the distribution of vorticity around the cylinder, while the change of drag can be concerned with the change of flow pressure. As shown in Fig. 13, there are two regions around the vortex-induced vibration cylinder, where pressure changes quite frequently. The first is the windward region of the front surface of the cylinder, and the second is the region near the separation point of the rear surface. Considering the suppression of the vibration, it can be explained that the pressure changes gently in the controlled windward region. The decrease in pressure drop gradient near the separation region behind the cylinder is due to the active flow control. When the average pressure of the two poles is almost the same, the DRL control reduces the reverse pressure gradient in the downstream of jet, which means that the pressure distribution of the cylinder wake is more well-distributed. The decrease in pressure difference between the front and back surfaces accounts for the drag reduction directly in Fig. 10.

C. Association between AL and DRL

Artificial neural network is an effective tool, which allows researchers to gain scientific achievements directly in disregard of the complex physical background. However, the artificial intelligence work lacks physical interpretability. For the periodic problem of VIV at Reynolds number 100, we will compare the control strategy obtained by active learning with that by DRL and seek a reason why the control strategy is successful and effective. As shown in Fig. 14, three control agents are selected in this comparison, including a small-action ($A = 0.8$, $\phi = 1.35$) triangle AL agent, a large-action ($A = 3.0$, $\phi = 1.35$) triangle AL agent, and a DRL agent. Among them, the agent from the DRL framework is the best agent, which successfully suppresses the vibration amplitude in a very small range. Compared with

the other two agents, the randomness of the action given by the DRL agent is hard to interpret, while there are still some similarities and associations between action and position worth discussing.

The triangle AL agent with large action presents the effect of large action on rapid vibration suppression, compared with small-action AL agent. Nevertheless, the continuous large-action control is not responsible for keeping the cylinder in a low amplitude vibration. On the contrary, it leads the vibration to a higher amplitude than the uncontrolled case after reaching the lowest point. Different from the large-action triangle AL agent, the small-action periodic control reduces the vibration amplitude of the cylinder to a certain degree and maintains such a vibration amplitude, which has very good stability and economic benefits. An apparent disadvantage of small-action AL agent is its limited control effect.

In the previous analysis of the instantaneous vorticity diagram, it is found that the strength of shedding vortices can be weakened with jet injection in the generation and suction in the shedding. The most direct embodiment of the above theory is that the action phase and position phase of the small-action periodic case are always opposite, which is also the contribution factor of the smooth control of the small action periodic function. The same is true for large action periodic function during the descent. With the amplitude decreasing, the phase of the displacement begins to change. When the vibration is stabilized again, a new in-phase vibration mode is formed between the action and position. The most direct explanation of the in-phase phenomenon is that the reaction force caused by the large action begins to directly drive the cylinder to vibrate at its frequency. This forced vibration is very similar to the action of slapping basketball, whose amplitude is related to the driving force. An antiphase is also found in the early amplitude decreasing stage directed by large-action AL agent, followed by a same-phase mode. What the DRL agent has done is to

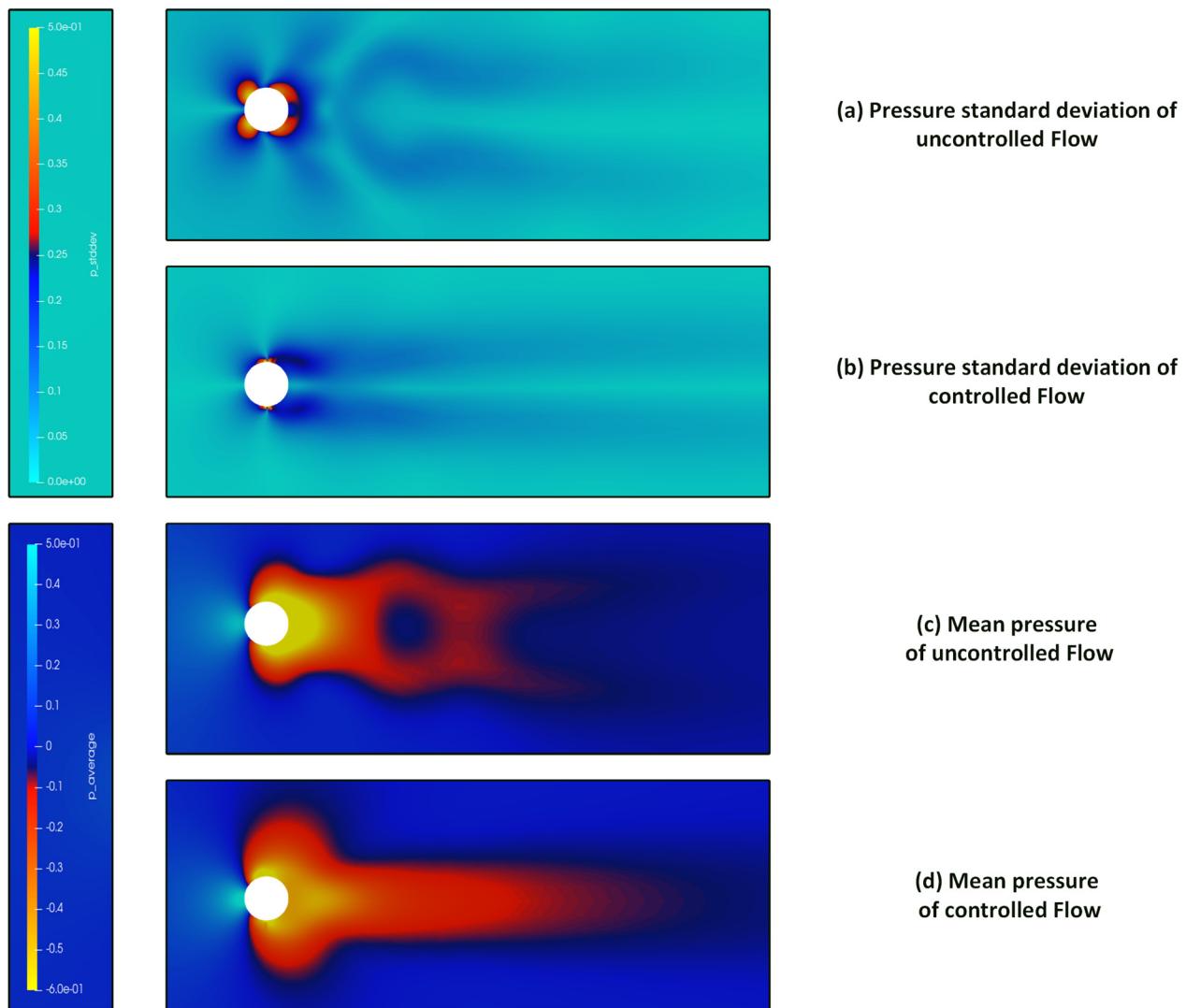


FIG. 13. Pressure temporal statistics contours. The Standard Deviation (STD) results in (a) and (b) point out that the separate region in the downstream of the control is the main battlefields of control. The comparison in mean pressure contours (c) and (d) shows the pressure difference is controlled, which explains the drag reduction directly.

provide a combination mode between large-action and small-action. The first step is to reduce the amplitude as quickly as possible with large action. The next operation is to adapt to the change of phase and amplitude of environment model by adjusting its own phase and amplitude. To stabilize the small amplitude vibration steadily, small-action strategy is carried out. To avoid a forced vibration mode, real-time adjustment is vital in the face of the change of phase caused by the large action, so that a state of reverse phase is maintained.

IV. CONCLUSION

In this study, an effective active flow control strategy is searched for to eliminate the vortex-induced vibration of the cylinder at low Reynolds number of $Re = 100$. The AFC based on a pair of symmetric jets is guided by the decision control agents, which are setup in the

frameworks of active learning and deep reinforcement learning. All the data are collected from the numerical simulation of computational fluid dynamics on OpenFOAM, which is integrated into a model named the environment. Every interaction between agent and environment is accompanied by control action and its corresponding flow state and dynamic feedback. Gaussian process regression method is used to establish a surrogate model in AL framework, which has an input of a control parameter set and an output of a stable vibration amplitude. With the aid of the Bayesian optimization method, it is able to reduce the amount of calculation while guaranteeing a high precision. Considering different properties of the open-loop control function, the triangle control agent has a natural advantage over the others in the periodic VIV problem. A key finding is that the small-action triangle AL agent stabilizes the vibration smoothly with an

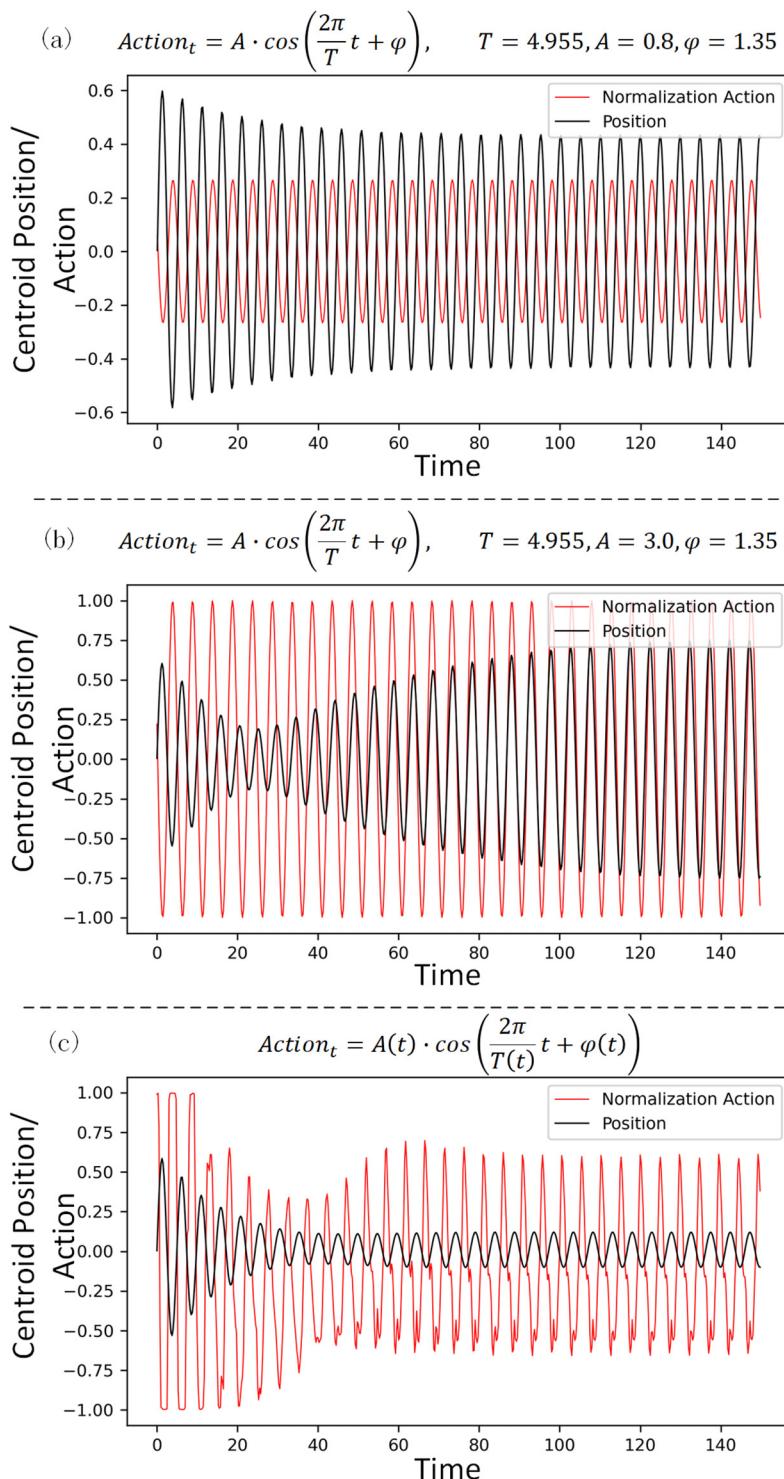


FIG. 14. Correlation between position trajectory and action trajectory. (a) Control action is given by small-action triangle agent. (b) Control action is given by large-action triangle agent. (c) Control action is given by DRL agent. Large action brings a faster rate of amplitude descent but keeps a high forced amplitude in the stable stage like bouncing basketball. Small action keeps track of the opposite phase between position and action, which makes contribution to a low-amplitude motion. The DRL agent combines the advantages of both that adopted large action control at first and adjust action amplitude or frequency according to current kinematics state throughout.

amplitude of 0.43, and the large-action triangle AL agent has the capability to suppress the vibration rapidly.

The DRL agent based on the soft actor-critic algorithm is a model-free agent, which has a requirement of a great quantity of

calculation. According to the results, the amplitude of the cylinder motion controlled by the DRL agent is reduced below 0.11, which decreased by 82.7% compared to the uncontrolled cylinder. What is behind the great success of the deep reinforcement learning algorithm

is the fact that the agent learns a key to the adjustment of the time-variant action amplitude and phase, which is hard to model and discuss in typical physical theory analysis. With large actions employed at first for speed, the small-action control is chosen ultimately to avoid a high-amplitude forced vibration. The reason to adopt an antiphase strategy between position and action is the injection during the vortex generation stage and the suction during the vortex shedding stage. The vorticity strength and distribution changed under active flow control, which leads to lower-level fluid-structure interactions. Meanwhile, a lower pressure drop is found near the separation region on the surface of the cylinder, which accounts for the drag reduction.

Despite the relative simplicity of the selected physical background, it is a brand-new attempt to introduce active learning and the deep reinforcement learning algorithm into the active flow control of the vortex-induced vibration problem. Without prior knowledge, it is hard to decide which form of active flow control in the AL framework is effective. However, the use of active learning makes a great contribution to the analysis of DRL control strategy and it shows a global optimal solution of visible pair in the specified action space, which saves the most computational resource and time. Deep reinforcement learning gains an extraordinary control effect at the cost of tremendous data, which brings a burden on practical application. In our research, numerical simulation takes up most of the time cost, compared with the training of artificial neural network. However, with the development of computational technology and data-driven engineering methods, such as, the reduced-order model, it seems promising to take deep reinforcement learning into industrial practice. On the basis of ensuring the calculation accuracy, how to reasonably configure the system state (sensor distribution⁴⁶), establish an appropriate reduced-order model,⁴⁷ and improve the calculation speed will be our next research objectives. The active learning and deep reinforcement learning are both popular machine learning technologies, which offer new choices in the exploration of active flow control. Similarly, there are many new machine learning methods^{48,49} worthy of being applied to unsteady VIV problem, where we are expecting to see more progress and possibilities.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support by the Natural Science Foundation of Zhejiang Province (Grant No. LY21A020010).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

- ¹V. Strouhal, "Über eine besondere art der Tonerregung," *Ann. Phys.* **241**, 216–251 (1878).
- ²F. Xie, D. Pan, Y. Zheng, and J. Zou, "Smoothed profile method and its applications in VIV," *Int. J. Numer. Methods Heat Fluid Flow* **27**, 1623–1635 (2017).
- ³T. Sarıkaya, "A critical review of the intrinsic nature of vortex-induced vibrations," *J. Fluids Struct.* **19**, 389–447 (2004).
- ⁴C. Williamson and R. Govardhan, "Vortex-induced vibrations," *Annu. Rev. Fluid Mech.* **36**, 413–455 (2004).
- ⁵A. Trim, H. Braaten, H. Lie, and M. Tognarelli, "Experimental investigation of vortex-induced vibration of long marine risers," *J. Fluids Struct.* **21**, 335–361 (2005).
- ⁶R. Galvao, E. Lee, D. Farrell, F. Hover, M. Triantafyllou, N. Kitney, and P. Beynet, "Flow control in flow–structure interaction," *J. Fluids Struct.* **24**, 1216–1226 (2008).
- ⁷F. Sher, S. Ahmed, and U. Asghar, "Suppression of vortex induced vibrations in marine risers by fairings," in *Proceedings of 14th International Bhurban Conference on Applied Sciences and Technology (IBCAST)* (IEEE, 2017), pp. 523–529.
- ⁸F. Xie, Y. Yu, Y. Constantinides, M. S. Triantafyllou, and G. E. Karniadakis, "U-shaped fairings suppress vortex-induced vibrations for cylinders in cross-flow," *J. Fluid Mech.* **782**, 300 (2015).
- ⁹W.-L. Chen, G.-B. Chen, F. Xu, Y.-w. Huang, D.-L. Gao, and H. Li, "Suppression of vortex-induced vibration of a circular cylinder by a passive-jet flow control," *J. Wind Eng. Ind. Aerodyn.* **199**, 104119 (2020).
- ¹⁰J. P. Pontaza and R. G. Menon, "Numerical simulations of flow past an aspirated fairing with three degree-of-freedom motion," in *Proceedings of International Conference on Offshore Mechanics and Arctic Engineering* (ASME, 2008), Vol. 48227, pp. 799–807.
- ¹¹Y. Delaunay and L. Kaitkis, "Control of circular cylinder wakes using base mass transpiration," *Phys. Fluids* **13**, 3285–3302 (2001).
- ¹²J. Kim and H. Choi, "Distributed forcing of flow over a circular cylinder," *Phys. Fluids* **17**, 033103 (2005).
- ¹³D. You and P. Moin, "Active control of flow separation over an airfoil using synthetic jets," *J. Fluids Struct.* **24**, 1349–1357 (2008).
- ¹⁴A. Barbagallo, G. Dergham, D. Sipp, P. J. Schmid, and J.-C. Robinet, "Closed-loop control of unsteadiness over a rounded backward-facing step," *J. Fluid Mech.* **703**, 326 (2012).
- ¹⁵N. Gautier, J.-L. Aider, T. Duriez, B. Noack, M. Segond, and M. Abel, "Closed-loop separation control using machine learning," *J. Fluid Mech.* **770**, 442–457 (2015).
- ¹⁶Y. Zhou, D. Fan, B. Zhang, R. Li, and B. R. Noack, "Artificial intelligence control of a turbulent jet," *arXiv:2005.04650* (2020).
- ¹⁷C. Lee, J. Kim, D. Babcock, and R. Goodman, "Application of neural networks to turbulence control for drag reduction," *Phys. Fluids* **9**, 1740–1747 (1997).
- ¹⁸F.-F. Xie, J. Deng, and Y. Zheng, "Multi-mode of vortex-induced vibration of a flexible circular cylinder," *J. Hydraul.* **23**, 483–490 (2011).
- ¹⁹S. Dong, G. Triantafyllou, and G. Karniadakis, "Elimination of vortex streets in bluff-body flows," *Phys. Rev. Lett.* **100**, 204501 (2008).
- ²⁰H. Zhu, T. Tang, H. Zhao, and Y. Gao, "Control of vortex-induced vibration of a circular cylinder using a pair of air jets at low Reynolds number," *Phys. Fluids* **31**, 043603 (2019).
- ²¹W.-L. Chen, D.-B. Xin, F. Xu, H. Li, J.-P. Ou, and H. Hu, "Suppression of vortex-induced vibration of a circular cylinder using suction-based flow control," *J. Fluids Struct.* **42**, 25–39 (2013).
- ²²D. R. Jones, M. Schonlau, and W. J. Welch, "Efficient global optimization of expensive black-box functions," *J. Global Optim.* **13**, 455–492 (1998).
- ²³D. Fan, G. Jodin, T. Consi, L. Bonfiglio, Y. Ma, L. Keyes, G. E. Karniadakis, and M. S. Triantafyllou, "A robotic intelligent towing tank for learning complex fluid-structure dynamics," *Sci. Rob.* **4**, eaay5063 (2019).
- ²⁴M. A. Gelbart, J. Snoek, and R. P. Adams, "Bayesian optimization with unknown constraints," *arXiv:1403.5607* (2014).
- ²⁵H. Zheng, F. Xie, T. Ji, Z. Zhu, and Y. Zheng, "Multifidelity kinematic parameter optimization of a flapping airfoil," *Phys. Rev. E* **101**, 013107 (2020).
- ²⁶C. Talnikar, P. Blonigan, J. Bodart, and Q. Wang, "Parallel optimization for LES," in *Proceedings of the Summer Program* (Stanford University, 2014), p. 315.
- ²⁷O. Mahfoze, A. Moody, A. Wynn, R. Whalley, and S. Laizet, "Reducing the skin-friction drag of a turbulent boundary-layer flow with low-amplitude wall-normal blowing within a Bayesian optimization framework," *Phys. Rev. Fluids* **4**, 094601 (2019).
- ²⁸V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, "Playing Atari with deep reinforcement learning," *arXiv:1312.5602* (2013).
- ²⁹S. Gu, E. Holly, T. Lillicrap, and S. Levine, "Deep reinforcement learning for robotic manipulation with asynchronous off-policy updates," in *Proceeding of 2017 IEEE International Conference on Robotics and Automation (ICRA)* (IEEE, 2017), pp. 3389–3396.
- ³⁰G. Reddy, A. Celani, T. J. Sejnowski, and M. Vergassola, "Learning to soar in turbulent environments," *Proc. Natl. Acad. Sci. U. S. A.* **113**, E4877–E4884 (2016).

- ³¹G. Reddy, J. Wong-Ng, A. Celani, T. J. Sejnowski, and M. Vergassola, "Glider soaring via reinforcement learning in the field," *Nature* **562**, 236–239 (2018).
- ³²S. Verma, G. Novati, and P. Koumoutsakos, "Efficient collective swimming by harnessing vortices through deep reinforcement learning," *Proc. Natl. Acad. Sci. U. S. A.* **115**, 5849–5854 (2018).
- ³³V. Belus, J. Rabault, J. Viquerat, Z. Che, E. Hachem, and U. Reglade, "Exploiting locality and translational invariance to design effective deep reinforcement learning control of the 1-dimensional unstable falling liquid film," *AIP Adv.* **9**, 125014 (2019).
- ³⁴J. Rabault, M. Kuchta, A. Jensen, U. Réglade, and N. Cerardi, "Artificial neural networks trained through deep reinforcement learning discover control strategies for active flow control," *J. Fluid Mech.* **865**, 281–302 (2019).
- ³⁵J. Rabault and A. Kuhnle, "Accelerating deep reinforcement learning strategies of flow control through a multi-environment approach," *Phys. Fluids* **31**, 094105 (2019).
- ³⁶H. Tang, J. Rabault, A. Kuhnle, Y. Wang, and T. Wang, "Robust active flow control over a range of Reynolds numbers using an artificial neural network trained through deep reinforcement learning," *Phys. Fluids* **32**, 053605 (2020).
- ³⁷R. Paris, S. Beneddine, and J. Dandois, "Robust flow control and optimal sensor placement using deep reinforcement learning," *J. Fluid Mech.* **913**, A25 (2021).
- ³⁸H. Jasak, A. Jemcov, Z. Tukovic *et al.*, "Openfoam: A c++ library for complex physics simulations," in *Proceedings of International Workshop on Coupled Methods in Numerical Dynamics* (IUC Dubrovnik, Croatia, 2007), Vol. 1000, pp. 1–20.
- ³⁹I. E. Barton, "Comparison of simple-and piso-type algorithms for transient flows," *Int. J. Numer. Methods Fluids* **26**, 459–483 (1998).
- ⁴⁰C. W. Hirt, A. A. Amsden, and J. Cook, "An arbitrary Lagrangian–Eulerian computing method for all flow speeds," *J. Comput. Phys.* **14**, 227–253 (1974).
- ⁴¹F. Xie, J. Deng, Q. Xiao, and Y. Zheng, "A numerical simulation of VIV on a flexible circular cylinder," *Fluid Dyn. Res.* **44**, 045508 (2012).
- ⁴²A. Wilson and H. Nickisch, "Kernel interpolation for scalable structured gaussian processes (KISS-GP)," in *Proceedings of International Conference on Machine Learning* (PMLR, 2015), pp. 1775–1784.
- ⁴³J. Nocedal, "Updating quasi-Newton matrices with limited storage," *Math. Comput.* **35**, 773–782 (1980).
- ⁴⁴T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor," in *Proceedings of International Conference on Machine Learning* (PMLR, 2018), pp. 1861–1870.
- ⁴⁵T. Haarnoja, A. Zhou, K. Hartikainen, G. Tucker, S. Ha, J. Tan, V. Kumar, H. Zhu, A. Gupta, P. Abbeel *et al.*, "Soft actor-critic algorithms and applications," *arXiv:1812.05905* (2018).
- ⁴⁶I. Akhtar, J. Borggaard, J. A. Burns, H. Imtiaz, and L. Zietsman, "Using functional gains for effective sensor location in flow control: A reduced-order modelling approach," *J. Fluid Mech.* **781**, 622 (2015).
- ⁴⁷I. Goumire, C. W. Rowley, Z. Ma, D. Gates, J. Krommes, and J. Parker, "Reduced-order model based feedback control of the modified Hasegawa–Wakatani model," *Phys. Plasmas* **20**, 042501 (2013).
- ⁴⁸H. Eivazi, H. Veisi, M. H. Naderi, and V. Esfahanian, "Deep neural networks for nonlinear model order reduction of unsteady flows," *Phys. Fluids* **32**, 105104 (2020).
- ⁴⁹T. Nakamura, K. Fukami, K. Hasegawa, Y. Nabae, and K. Fukagata, "Convolutional neural network and long short-term memory based reduced order surrogate for minimal turbulent channel flow," *Phys. Fluids* **33**, 025116 (2021).